Jewgeni H. Dshalalow

Foundations of Abstract Analysis

Second Edition

Springer

Foundations of Abstract Analysis

Jewgeni H. Dshalalow

Foundations of Abstract Analysis

Second Edition

Springer

Jewgeni H. Dshalalow
Mathematical Sciences
Florida Institute of Technology
Melbourne 32901
Florida
USA

ISBN 978-1-4614-5961-3 ISBN 978-1-4614-5962-0 (eBook)
DOI 10.1007/978-1-4614-5962-0
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012951666

Mathematics Subject Classification (2010): 54-01, 28A05, 28A10, 28A12, 28A20, 28A25, 28A33, 28A35, 26-XX, 26-01

2nd Ed of a book previously published by Chapman and Hall/CRC Press in 2000, *Real Analysis: An Introduction to the Theory of Real Functions and Integration*

© Springer Science+Business Media, LLC 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Irina

Preface

This is the first of a two book series written for beginning graduate-level real analysis students and it is focused on essentials of set theory, topology, and measure theory. The word "essentials" is often associated with lecture notes, but this is exactly opposite to my intention. While there are quite a few fine abstract analysis books, they are difficult to read, and an enthusiastic reader needs more time and effort to master the subject, as opposed to elementary level books that only in part address the needs of forthcoming mathematics and advanced engineering courses.

That is why this book offers a thorough and yet rigorous treatment of key analysis subjects in abstract spaces by providing the reader with copious illustrations, examples, and exercises with selected solutions. Some difficult to understand topics are preceded by detailed discussions and blueprints, so that the reader will not get lost or intimidated in a long chain of proofs and notions. Furthermore, at the end of each section there is a summary of new terms and notation in the chronological order in which they appeared in the text. This should be helpful and relieve students of the potential overload of new words, definitions, and concepts. So, there is an increased likelihood of success by using this text in a course or by means of individual self-study.

My previous text, *Real Analysis*, published by Chapman and Hall in 2000, was the first effort to create this kind of book. However, this book only partially accomplished the goal I was striving to achieve. To fully realize that goal, it was necessary to write a new and expanded edition, including more topics and details, and it had to be produced as two books. The companion book, *Advances in Abstract Analysis and Applications*, includes further topics in topology and

measure theory, which justifies and rewards the reader for investing the time spent on "essentials."

As mathematical education has become increasingly more focused on applications and less on theory, and in order to save them from extinction, academics have repurposed courses in set theory, topology, abstract algebra, and measure and integration as a *real analysis* course. At the same time, mathematical research, driven by serious applications to other sciences, continued to require sound foundations. The pertinent precedents include physics, stochastic finance, mechanics, and now biology. There was even a time when some proponents called real analysis "the single most important graduate course in mathematics to prepare for a career in operations research."

Today, real analysis is still very much alive, although it has undergone some significant modifications. One of these changes is that contemporary real analysis books include various, sometimes exotic, applications ranging from partial differentials equations to wavelet analysis, probability, and even physics. While such connections might be justified, one has to ensure that this propensity to connect analysis with remote disciplines does not relax its very substance. Consequently, facing a challenge of two alternatives to yield an oversized or abridged book, a broad-spectrum project (under strong encouragement from Springer) got bifurcated into two entirely differently focused texts.

Because real analysis, in its proper form, is likely to be the first abstract mathematics course that many students take, the associated topics should be taught in a strict order starting with basic set theory followed by point-set topology and then measure theory and integration. Throughout my book I follow these principles. I strongly advocate the idea of introducing measure and integration in abstract spaces wasting no valuable time on Euclidean spaces. Consequently, Lebesgue measure and Lebesgue integral are reduced to mere illustrations. The topology part, to be necessarily preceded by metric spaces, contains mostly fundamentals (such as bases, subbases, Hausdorff spaces, Tychonov product, and compactness). In particular, old good sequential convergence is enough to proceed with rigorous and comprehensive measure and integration (which accounts

to over 55% of my book). Such topics as filters and nets, locally compact Hausdorff spaces, Radon measures, and Hilbert spaces (measure-theoretic version) I consider relatively advanced and therefore treat them in my forthcoming sequel to this book, along with various applications to stochastic analysis.

It is absolutely impossible to produce a sound text without building on the foundations of my predecessors' important scholarship. I am very grateful for the valuable remarks and suggestions, made to the present and earlier edition, by Gustave Choquet, Jerald Folland, Jordan Stoyanov, Jürgen Becker, Richard Syski, Jean Lasserre, Donald Konwinski, and Dean Spitzer. I am much indebted to Simon Smith, the creator of the EXP word processor, for his generous and timely support.

My deep appreciation also goes to Ms. Vaishali Damle, who is an extremely competent and efficient mathematics editor. Her constant help and patience was made available from the very beginning of the project and receiving her feedback was one of the key benefits of writing the book. Finally, I want to thank the Springer copyeditor, Ms. Valerie Greco, who did an excellent job in polishing the final draft of my manuscript. Because books are often written at home, I would be remiss in not ending with profound gratitude to my wife, Irina, who created an environment conducive to writing this book.

J.H. Dshalalow
Melbourne, Florida

Contents

Preface *vii*

PART I
Principles of Set Theory and Point-Set Topology **1**

Chapter 1 Set-Theoretic and Algebraic Preliminaries **3**
1. Sets and Basic Notation 3
2. Functions 14
3. Set Operations under Maps 23
4. Relations and Well-Ordering Principle 29
5. Cartesian Product 42
6. Cardinality 58
7. Basic Algebraic Structures 66
8. More about Vector Spaces 85

Chapter 2 Analysis of Metric Spaces **93**
1. Definitions and Notations 94
2. The Structure of Metric Spaces 101
3. Convergence in Metric Spaces 112
4. Continuous Mappings in Metric Spaces 117
5. Complete Metric Spaces 132
6. Compactness 138
7. Normed Vector Spaces 149
8. Finite-Dimensional Normed Vector Spaces 160

Chapter 3 Elements of Point-Set Topology **171**
1. Topological Spaces 171
2. Bases and Subbases for Topological Spaces 188

3. Hausdorff Topological Spaces 197
4. Continuity in Topological Spaces 203
5. Product Topology 213
6. Notes on Subspaces and Compactness 226

PART II
Measure Theory and Abstract Integration **237**

Chapter 4 Measurable Spaces and Measurable Functions 239
1. Systems of Sets 240
2. System's Generators 247
3. Measurable Functions 255

Chapter 5 Measures 261
1. Set Functions 263
2. Extension of Set Functions to a Measure 283
3. Lebesgue and Lebesgue-Stieltjes Measures 313
4. Image Measures 336
5. Extended Real-Valued Measurable Functions 343
6. Simple Functions 354

Chapter 6 Integration in Abstract Spaces 361
1. Integration on $\mathcal{C}^{-1}(\Omega, \Sigma)$ 362
2. Main Convergence Theorems 382
3. Lebesgue and Riemann Integrals on $\mathbb{R}$ 399
4. Integration with Respect to Image Measures 416
5. Positive Measures Generated by Integrals. Absolute Continuity 422
6. Signed Measures 434
7. Complex Measures and Integration 458
8. Radon-Nikodym Theorem for Signed and Complex Measures and its Applications 471
9. Measures and Integration on Finite Product Spaces 498

PART III
Supplementary Materials **533**

Chapter 7 Solutions to Selected Problems **535**

Bibliography **715**
List of Symbols **721**

Index of Names and Terms **731**

Chapter 1

Set-Theoretic and Algebraic Preliminaries

Set theory is not just one of the main tools in mathematics, it is the very root of mathematics from which all mathematical disciplines stem. Due to a series of his papers, the first of which appeared in 1874, the great German mathematician, Georg Ferdinand Cantor (1845 – 1918), is considered to be the sole founder of set theory . Although the Czech Bernard Bolzano (1781 – 1848) made one of the first attempts to formalize set theory, in particular in his *Paradoxien des Unendlichen* (of 1851), by considering the one-to-one correspondence between two sets (later on developed by Cantor to what we now know as *cardinals*), neither he, nor anyone else, was really a predecessor to Cantor's creation. Ernst Zermelo (1871 – 1953) was another German, who among his numerous contributions to set theory, is the author of the first axiom for set theory (of 1908) and undoubtedly one of the primary axioms of the whole mathematics.

This chapter presents only essentials of set theory and abstract algebra needed throughout the book.

1. SETS AND BASIC NOTATION

Cantor defined a *set* as a *collection M into a whole of definite and distinct objects* (that are called *elements* of M) *of our thought.* In other words, we bind objects (perhaps of different nature) in our mind into a single entity and call that entity a *set.*

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0_1, © Springer Science+Business Media, LLC 2013

We denote sets by capital letters, and their elements by lower case letters. For instance, a set A has elements a, b, c, or $a_1, a_2, \ldots$. To abbreviate the expression "a is an element of the set A," we write $a \in A$. The expression "$a \notin A$" reads "a is not an element of A."

Observe that the notion of a set is relatively simple if we deal with such frequently encountered sets as sets of integers, rational numbers, real numbers, or continuous functions. In some rare situations, thoughtless use of this notion can lead to contradictions, as with Bertrand Russell's paradox. Russell posed the following set dilemma. Let $\mathcal{R}$ be the set of all sets, which are not elements of themselves. Clearly, $\mathcal{R}$ is not empty. For instance, the set of all real numbers is not an element of itself (for it is not a real number), thus it belongs to $\mathcal{R}$. The question arises: is $\mathcal{R}$ an element of itself? If $\mathcal{R} \in \mathcal{R}$ then by definition of $\mathcal{R}$, it should not belong to $\mathcal{R}$, which is a contradiction. Thus, $\mathcal{R} \notin \mathcal{R}$. But then, by definition, it must belong to $\mathcal{R}$, which is impossible. In this case, we have put the definition of an object ahead of its existence. The concept of a set must be supported by axioms of set theory, just as the main axioms of plane geometry define the shape of lines.

1.1 Definitions.

(i) A set A is said to be a *subset* of a set B (in notation, $A \subseteq B$) if all elements of A are also elements of B. If A is a subset of B, we call B a *superset of* A (in notation, $B \supseteq A$). A set that contains exactly one element, say a, is called a *singleton* (*set*) and it is denoted by $\{a\}$. If $a \in A$, then we can alternatively write $\{a\} \subseteq A$. Any set is obviously a subset of itself: $A \subseteq A$.

(ii) The unique set with no elements is called the *empty set* and is denoted Ø. Clearly, Ø is a subset of any set, including itself.

(iii) $A = B$ (read "set A equals set B") if and only if $A \subseteq B$ and $B \subseteq A$; otherwise, we will write $A \neq B$. Occasionally, we use the symbol " $\subset$ " to appropriate the situation where one set is a subset of another set but the sets are not equal. $A \subset B$ reads "A is a *proper subset* of B." In this case, B is a *proper superset of* A (in notation, $B \supset A$). □

We postulate the existence of a set that is a superset of all other sets in the framework of a certain mathematical model. This set is usually called a *universal set* or just *universe*. We also make use of the word "carrier" as a synonym for the universe and reserve for it the Greek letter Ω. Sometimes, we denote it by X, Y, or Z. A universe (as a base for some mathematical model or problem) is generally defined to contain all considered sets and it varies from model to model. For example, if $C^n_{[a,b]}$ denotes the set of all n-times differentiable functions on interval $[a,b]$, it contains, as a subset, the set of possible solutions of an ordinary differential equation of the nth order. Thus, $\Omega = C^n_{[a,b]}$ is a relevant universe within which the problem is posed. One could also take for Ω the set $C_{[a,b]}$ of all continuous functions on $[a,b]$ or even the set of all real-valued functions on $[a,b]$. However, these are "vast" to serve for universes and they are impractical for this concrete problem.

Set theory is also a basic component of probability theory, which always begins with elements of set theory under a slightly modified lexicon. For instance, a universe is referred to as a *sample space.* Specially chosen (or *measurable*) subsets of the sample space are called *events.* In particular, measurable singletons are called *elementary events.*

The concept of the universe becomes most vivid when used in probability theory. Let us consider the experiment that consists of tossing a coin until the first appearance of the head on the upper face of the coin. Denoting H as an output of the head and T as an output of the tail, when tossing the coin, we may define $\{(T, T, \dots, T, H)\}$ as an elementary event of the sample space Ω filled with the elements $\{(H), (T, H), (T, T, H), \dots\}$. The sample space Ω contains, as elements, all possible outcomes of tossing the coin until the "first success" or the first appearance of the head. For instance, the event $\{(H), (T, H), (T, T, H)\}$ corresponds to the "success in at most three tosses."

1.2 Notations. Throughout the whole book we use the following notation.

(i) Logical symbols:

$\forall$ means *"for all"*.
$\exists$ means "*there is*" or "*there are*" or "*there exists*".
$\Rightarrow$ means "*implies*" or "*from ... it follows that ...*".
$\Leftrightarrow$ means "*if and only if*".
$\wedge$ means "*and*" (also "&" on occasion).
$\vee$ means "*or*".
: means "*such that*" (used for definition of sets).

(ii) Frequently used sets:

$\mathbb{N}$: the set of all positive integers.
$\mathbb{N}_0$: the set of all nonnegative integers.
$\mathbb{Z}$: the set of all integers.
$\mathbb{Q}$: the set of all rational numbers.
$\mathbb{Q}^c$: the set of all irrational numbers.
$\mathbb{R}$: the set of all real numbers.
$\mathbb{C}$: the set of all complex numbers.
$\mathbb{R}_+$: the set of all nonnegative real numbers.
$\mathbb{R}_-$: the set of all negative real numbers.

(iii) Denotation of sets:

List: The elements are listed inside a pair of braces [for instance, $\{a, b, c\}$ or $\{a_1, a_2, \dots\}$].
Condition: A description of the elements with a condition following a colon (that in this case reads "such that"), again with braces enclosing the set (for instance, *The set of odd integers* is $\{n \in \mathbb{Z}: n = 2k + 1, k \in \mathbb{Z}\}$).

(iv) Main set operations:

Union: $A \cup B = \{x \in \Omega: x \in A \vee x \in B\}$.
Intersection: $A \cap B = \{x \in \Omega: x \in A \wedge x \in B\}$.
Two subsets $A, B \subseteq \Omega$ are called *disjoint* if $A \cap B = \emptyset$.
Difference: $A \backslash B = \{x \in \Omega: x \in A \wedge x \notin B\}$ ($A \backslash B$ is also called the *complement of B with respect to A*, with the alternative notation $A - B$ or B^c_A.)

Symmetric Difference: $A \triangle B = (A \backslash B) \cup (B \backslash A)$.
Complement (*with respect to the universe* Ω) : $A^c = A^c_\Omega = \Omega \backslash A$.

(v) *General notation*:

"$:=$" reads "*set by definition.*"
$\square$ means the end of a proof, remarks, examples, and so on.

A *set-algebraic expression* is a set in the form of some defined sets connected through set operations. Any transformation of a set-algebraic expression into another expression would require a set-theoretic manipulation we call a *set-algebraic transformation.* All basic set-algebraic transformations over basic set-algebraic expressions are known as *laws of algebra* (or *calculus*) *of sets.* $\square$

1.3 Remark. One of the standard tools of the algebra of sets is the so-called *pick-a-point process* applied to, say, showing that $A \subseteq B$ or $A = B$. It is based on the following.

Axiom of Extent*: For each set* A *and each set* B, *it is true that* $A = B$ *if and only if for every* $x \in \Omega$, $x \in A$ *when and only when* $x \in B$.

Axiom's modification:

If every element of A *is an element of* B, *then* $A \subseteq B$.

Thus, for the modification, the pick-a-point process consists of selecting an arbitrary point x of A (picking a point x) and then proving that x also belongs to B.

The identities below can be verified easily by the reader using pick-a-point techniques. $\square$

1.4 Theorem (Laws of Algebra of Sets).

(i) *Commutative Laws*:
$A \cup B = B \cup A$.
$A \cap B = B \cap A$.

(ii) *Associative Laws*:
$(A \cup B) \cup C = A \cup (B \cup C)$.

$$(A \cap B) \cap C = A \cap (B \cap C).$$

(iii) *Distributive Laws*:
$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$
$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$

(iv) *Idempotence of*
complement: $(A^c)^c = A$
union: $A \cup A = A$
intersection: $A \cap A = A.$

(v) $A \cap A^c = \emptyset.$

(vi) $A \cup A^c = \Omega.$

(vii) *De Morgan's Laws*:
$(A \cup B)^c = A^c \cap B^c.$
$(A \cap B)^c = A^c \cup B^c.$

$(viii)$ $A \cup \emptyset = A.$

(ix) $A \cap \emptyset = \emptyset.$

(x) $\Omega^c = \emptyset$ and $\emptyset^c = \Omega.$ □

1.5 Example. Show the validity of the first distributive law:

$$\begin{aligned}
& x \in (A \cup B) \cap C \\
\Leftrightarrow\ & x \in (A \cup B) \ \wedge\ x \in C \\
\Leftrightarrow\ & [x \in A \ \wedge\ x \in C \\
\Leftrightarrow\ & x \in A \cap C] \ \vee\ [x \in B \ \wedge\ x \in C \ \Leftrightarrow\ x \in B \cap C] \\
\Leftrightarrow\ & x \in (A \cap C) \cup (B \cap C).
\end{aligned}$$ □

1.6 Remark. The concepts of union and intersection can be extended to an arbitrary family of sets. For instance,

$$\bigcup_{i \in I} A_i = \{x \in \Omega : \exists\, i \in I,\ x \in A_i\}.$$

The distributive laws and De Morgan's laws hold for arbitrary families [subject to Problem 1.1b)]:

$$\left(\bigcup_{i\in I} A_i\right) \cap B = \bigcup_{i\in I} (A_i \cap B)$$

$$\left(\bigcap_{i\in I} A_i\right) \cup B = \bigcap_{i\in I} (A_i \cup B)$$

$$\left(\bigcup_{i\in I} A_i\right)^c = \bigcap_{i\in I} A_i^c$$

$$\left(\bigcap_{i\in I} A_i\right)^c = \bigcup_{i\in I} A_i^c\,. \qquad \square$$

1.7 Definitions.

(i) An indexed family $\mathcal{F} = \{A_i \subseteq \Omega : i \in I\}$ of sets is called (*pairwise*) *disjoint*, if for all $i \neq j$, $A_i \cap A_j = \emptyset$.

Throughout this book, the union of a pairwise disjoint family of sets is denoted for convenience by $\sum_{i\in I} A_i$. Specifically, $A + B$ means $A \cup B$, when A and B are disjoint.

(ii) A *decomposition* of a set A is any representation of A as the union of a disjoint family of sets, $A = \sum_{i\in I} A_i$. The family $\{A_i; i \in I\}$ is referred to as a *partition* of A. (There is another use of the term *partition*, applied to a different construction in a narrower sense. Namely, P is a *partition* of a closed interval $[a, b] \subseteq \mathbb{R}$ if P is any ordered finite set of points $\{a_0, \ldots, a_n\} \subseteq [a, b]$ with $a = a_0 < a_1 < \ldots < a_n = b$.)

(iii) Let Ω be a fixed set. The family of all subsets of Ω is called the *power set of* Ω and it is denoted by $\mathcal{P}(\Omega)$.

(iv) A sequence $\{A_n : n = 1, 2, \ldots\}$ of sets is said to be *monotone nondecreasing* (*nonincreasing*), if

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq \cdots (A_1 \supseteq A_2 \supseteq \cdots \supseteq A_n \supseteq \cdots).$$

To specify the type of convergence, we write $\{A_n\} \uparrow A$ ($\{A_n\} \downarrow A$). A sequence $\{A_n\}$ of sets is said to be *monotone vanishing*, if it is monotone nonincreasing and $\{A_n\} \downarrow \emptyset$.

(v) Let $\{A_n\}$ be an arbitrary sequence of sets. Denote

(a) $\liminf_{n\to\infty} A_n$ (or just $\underline{\lim}\ A_n$) $= \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$. This limit is called the *limit inferior.*

(b) $\limsup_{n\to\infty} A_n$ (or just $\overline{\lim}\ A_n$) $= \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$. This limit is called the *limit superior.*

If $\underline{\lim}\ A_n = \overline{\lim}\ A_n$, then we denote their common limit by $\lim_{n\to\infty} A_n$. In this case, the *limit of* (the sequence) $\{A_n\}$ is said to *exist* and equal $\lim_{n\to\infty} A_n$. □

PROBLEMS

1.1 a) Prove Theorem 1.4, the laws of algebra of sets by using the pick-a-point process.
b) Prove the generalized distributive laws and De Morgan's laws stated in Remark 1.6.

1.2 Show that:

a) $(A \cup B)\backslash C = (A\backslash C) \cup (B\backslash C)$.
b) $(A \cap B)\backslash C = (A\backslash C) \cap (B\backslash C)$.
c) $C\backslash(A \cup B) = (C\backslash A) \cap (C\backslash B)$.
d) $C\backslash(A \cap B) = (C\backslash A) \cup (C\backslash B)$.
e) $(A\backslash B) \cap C = A \cap (C\backslash B)$.

1.3 Show that $A\backslash B = A \cap B^c$.

1.4 Let $|A| = n$ (i.e., the set A contains n elements). Show that $|\mathcal{P}(A)| = 2^n$.

1.5 Prove that:

a) $(A\backslash B)^c = A^c \cup B$.
b) $[(A^c \cup B)^c \cup (A \cup B^c)]^c = B\backslash A$.

$c)\ (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B) = A \cup B\,.$

1.6 For each of the following, justify with a proof or give a counterexample.

$a)\ A \cup C = B \cup C \ \Rightarrow\ A = B.$
$b)\ (A \cup B) \backslash B = A.$
$c)\ A \backslash B = C \backslash B \ \Rightarrow\ A = C.$
$d)\ A^c \backslash B^c = B \backslash A.$

1.7 Give an example of a monotone vanishing sequence of sets.

1.8 Let $\{A_n :\ n = 1, 2, \dots\ \}$ be an arbitrary sequence of sets. Define $A_0 = \bigcap_{n=1}^{\infty} A_n$ and $A_\infty = \bigcup_{n=1}^{\infty} A_n$.

a) Construct a monotone nonincreasing sequence of sets $\{B_n\}$ such that $\{B_n\} \downarrow A_0$.

b) Construct a monotone nondecreasing sequence of sets $\{C_n\}$ such that $\{C_n\} \uparrow A_\infty$.

c) Given $\{C_n\} \uparrow A_\infty$, construct a pairwise disjoint sequence $\{D_n\}$ such that $\sum_{n=1}^{\infty} D_n = A_\infty$.

1.9 Let $\{A_n :\ n = 1, 2, \dots\}$ be an arbitrary sequence of sets. Define $A_0 = \bigcap_{n=1}^{\infty} A_n$ and $A_\infty = \bigcup_{n=1}^{\infty} A_n$. Prove that

$$A_0 \subseteq \underline{\lim} A_n \subseteq \overline{\lim} A_n \subseteq A_\infty.$$

1.10 Let Ω be an arbitrary set. Find a sequence $\{E_n\}$ of subsets of Ω such that

$$\underline{\lim}\ E_n = \emptyset \quad \text{and} \quad \overline{\lim}\ E_n = \Omega.$$

NEW TERMS:

Cantor, Georg, 3
Bolzano, Bernard, 3
Zermelo, Ernst, 3
set, 3
element of a set, 3
Russell's paradox, 4
subset, 4
superset, 4
singleton, 4
empty set, 4
proper subset, 4
proper superset, 4
universe, 5
carrier, 5
sample space, 5
measurable subset, 5
event, 5
elementary event, 5
Logical Symbols, 6
- $\forall$ means "for all"
- $\exists$ means "there is" or "there are" or "there exists"
- $\Rightarrow$ means "implies" or "from ... it follows that ..."
- $\Leftrightarrow$ means "if and only if"
- $\wedge$ means "and"
- $\vee$ means "or"
- : means "such that" (used for definition of sets)

Sets, 6
- $\mathbb{N}$: the set of all positive integers
- $\mathbb{N}_0$: the set of all nonnegative integers
- $\mathbb{Z}$: the set of all integers
- $\mathbb{Q}$: the set of all rational numbers
- $\mathbb{Q}^c$: the set of all irrational numbers
- $\mathbb{R}$: the set of all real numbers
- $\mathbb{C}$: the set of all complex numbers
- $\mathbb{R}_+$: the set of all nonnegative real numbers

$\mathbb{R}_-$: the set of all negative real numbers
union, 6
intersection, 6
disjoint sets, 6
difference, 6
complement of set B with respect to set A, 6
symmetric difference, 7
complement, 7
set-algebraic expression, 7
set-algebraic transformation, 7
laws of algebra of sets, 7
pick-a-point process, 7
axiom of extent, 7
commutative laws of sets, 7
associative laws of sets, 7
distributive laws of sets, 8
idempotence, 8
idempotence of complement, 8
idempotence of union, 8
idempotence of intersection, 8
De Morgan's laws, 8
pairwise disjoint sets, 9
disjoint family of sets, 9
decomposition of a set, 9
partition of a set, 9
partition of an interval, 9
power set, 9
monotone nondecreasing sequence of sets, 9
monotone nonincreasing sequence of sets, 9
monotone vanishing sequence of sets, 9
limit inferior of a sequence of sets, 10
limit superior of a sequence of sets, 10
limit of a sequence, 10

2. FUNCTIONS

The word "function" was introduced by Gottfried von Leibnitz in 1694, initially as a term to denote any quantity related to a curve, such as its slope, the radius of curvature, and so on. The notion of the function was refined subsequently by Johann Bernoulli, Leonard Euler, Joseph Fourier, and finally, by Gustav Dirichlet in the middle of the nineteenth century with a formulation pretty close to what we are using at the present time and which a mathematics or engineering student meets in an introductory calculus course. Dirichlet introduced a *variable*, as a symbol that represents a set of numbers; if two variables x and y are so related that whenever x takes on a value, there is a value y assigned to x by some rule of correspondence. In this case y (a *dependent variable*) was said to be a *function of* x (an *independent variable*).

In this section we introduce a more contemporary notion of a function. For functions operating with sets (rather than with points), we use a nontraditional notation of f_* and f^* (instead of just f), which we found very appealing as it brings more order within functions acting on collections of sets (such as topologies and sigma-algebras) and simplifies many proofs.

2.1 Definitions.

(i) Let X and Y be two sets. The set $\{(x, y)\colon\ x \in X,\ y \in Y\}$ of all ordered pairs of elements of X and Y is called the *Cartesian* or *direct product of* X *and* Y and it is denoted by $X \times Y$. If $X = Y$ then we write $X \times X = X^2$. Similarly, the *Cartesian product* of n sets,

$$X_1 \times \cdots \times X_n = \{(x_1, \ldots, x_n)\colon\ x_1 \in X_1, \ldots,\ x_n \in X_n\},$$

is the set of all ordered n-tuples.

(ii) Any subset f of $X \times Y$ is called a *binary relation.*

(iii) A binary relation $f \subseteq X \times Y$ is called a *single-valued function* (or just a *function*) if whenever (x, y_1) and (x, y_2) are elements of f, then $y_1 = y_2$. We also say that the function f is a

map (or *mapping* or *transformation*) *from* X *to* Y and denote this most frequently by the triple $[X, Y, f]$ or by $f: X \to Y$ or by $(x, f(x))$ or by $y = f(x)$ or by $x \mapsto f(x)$.

Note that, although both X and Y are components of triple $[X, Y, f]$, some pairs (x, y) from $X \times Y$ need not belong to f whether f is a function or a binary relation.

(iv) For a function f (as a subset of $X \times Y$), denote

$$D_f = \{x \in X: (x, y) \in f\}$$

and call it the *domain* of f. When a function $[X, Y, f]$ is given, its domain is generally not specified. The latter requires a more rigorous motivation. For instance, let $X = Y = \mathbb{R}$ and $f(x) = \sqrt{x^2 - 1}$. Then, $D_f = (-\infty, -1] \cup [1, \infty)$ and for any $x \in (-1, 1)$, there is no mate from Y under f. We can say that f is defined only for $x \in D_f$.

In another example, let $f(x) = \frac{1}{\sqrt{x-1}}$. This function is defined for all $x \in (1, \infty)$ if $X = Y = \mathbb{R}$. However, if Y is the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$, we allow $x \in [1, \infty]$. And finally, it is not wrong to have x be any real (or even complex) number, if f will take on values in $Y \subseteq \mathbb{C}$ (or $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$).

(v) Another component of a function is its *range*,

$$R_f = \{y \in Y: \exists\, x \in D_f\,,\ f(x) = y\}.$$

A superset of R_f (such as Y) is referred to as a *codomain*. In other words, R_f is the subset of all such elements of Y, which take part in the relation $f \subseteq D_f \times Y$.

(vi) If $x \in D_f$, then $f(x)$ $(\in R_f)$ is called the *image of* x *under* f. By the above definition of a single-valued function, for every x, there is a unique image $f(x)$. (An "extended" concept of a function allows more than one image of each point x under f. Any such function f is called *multivalued*. The reader is definitely acquainted with principles of complex analysis where such functions are

common. It is also known that in this case the range of a multi-valued function can be partitioned into pairwise disjoint subsets, such that the function is then split into a number of single-valued functions called *branches*.)

Note that in the primary definition $[X, Y, f]$ of a function most often Y is a codomain and X is a domain of f. However, in some cases, X is not a domain of f, which means that for some $x \in X$, there is no mate $y \in Y$ such that $(x, y) \in f$. In the latter case we call $[X, Y, f]$ *inconsistent*. For instance, in the above example, the function $\left[\mathbb{R}, \mathbb{R}, \sqrt{x^2 - 1}\right]$ is inconsistent, because there are no mates $y \in Y$ for all $x \in (-1, 1)$. If we try to fix this problem by a minimal invasion, namely assigning the empty set Ø to any such point (i.e., setting $f(x) = \text{Ø}$), we are up against an inconsistency, for $f(x)$ is suppose to be a point, not a set. The following construction of f_* better deals with this predicament. A function $[X, Y, f]$ is *consistent* if $X = D_f$.

(vii) If $A \subseteq D_f$ then the set of the images of all points of A under f is called the *image of A under f* and, following the notation of most analysis textbooks, it can be denoted

$$f(A) = \{y \in Y \colon \exists\, x \in A,\ f(x) = y\}.$$

However, for the upcoming constructions, it is convenient to distinguish images of points from images of subsets of X under f. In other words, we introduce the function $[\mathcal{P}(X), \mathcal{P}(Y), f_*]$, where for $A \in \mathcal{P}(X)$ we denote

$$f_*(A) = \{y \in Y \colon \exists\, x \in A,\ f(x) = y\}$$

Specifically, $R_f = f_*(D_f)$. In reference to inconsistency in (vi), we agree to set $f_*(\{x\}) = \text{Ø}\ \ \forall x \notin D_f$ or more generally, $f_*(A) = \text{Ø}$ whenever $A \notin \mathcal{P}(D_f)$. With this add-on, $[\mathcal{P}(X), \mathcal{P}(Y), f_*]$ becomes consistent. The latter alone makes f_* superior to f.

In some forthcoming applications, we also use the notion of function f_{**} defined on $\mathcal{P}(\mathcal{P}(X))$ and with $\mathcal{P}(\mathcal{P}(Y))$ as its codomain.

$(viii)$ Let $[X, Y, f]$ be a function. In its "ideal" construction call $[\mathcal{P}(R_f), \mathcal{P}(D_f), f^*]$ the *inverse of* f_*. In other words, for each set $B \in \mathcal{P}(R_f)$, the inverse image $f^*(B) = \{x \in X: \ f(x) \in B\}$. The set $f^*(B)$ is called the *inverse image of* B *under* f, or the *pre-image of* B *under* f.

In a more general setting we define the *inverse of* f_* as $[\mathcal{P}(Y), \mathcal{P}(X), f^*]$ bearing in mind that $f^*(B) = \emptyset$ if $B \notin \mathcal{P}(R_f)$.

Another construction related to f^* and often used in analysis is f^{-1} defined as $\{(y, x) \in Y \times X : (x, y) \in f\}$ and called the *inverse of* f. Unlike f^*, f^{-1} need not be a single-valued function (in other words, it is a binary relation or multivalued function). Consider, for instance, the function $[\mathbb{R}, \mathbb{R}, f]$ such that $f(x) = x^2$. Clearly, $R_f = \mathbb{R}_+$ and the inverse $\sqrt{} = f^{-1}$ of f is a two-valued function with domain $D_{f^{-1}} = \mathbb{R}_+$ and with range equal $\mathbb{R}$, which can be decomposed as $\mathbb{R} = (-\infty, 0) + [0, \infty)$. Accordingly, we have two branches $[\mathbb{R}_+, (-\infty, 0), \sqrt{}\,]$ and $[\mathbb{R}_+, \mathbb{R}_+, \sqrt{}\,]$ of $\sqrt{}$.

In light of the function $[\mathcal{P}(\mathcal{P}(X)), \mathcal{P}(\mathcal{P}(Y)), f_{**}]$ we likewise consider its inverse $[\mathcal{P}(\mathcal{P}(R_f)), \mathcal{P}(\mathcal{P}(X)), f^{**}]$.

(ix) Observe that it is legitimate for a single-valued function to have $f(x_1) = f(x_2)$ and $x_1 \neq x_2$. However, if f is such that $f(x_1) = f(x_2)$ if and only if $x_1 = x_2$, then f is called *one-to-one* or *injective* or *injection*. If f is injective, f^{-1} is a single-valued function too. For example, the map $[\mathcal{P}(R_f), \mathcal{P}(D_f), f^*]$ is injective. (See Problem 2.8.)

Because f^{-1} in general is not a single-valued function we agree to regard $f^{-1}(y)$ as a set (which in particular can be a singleton or the empty set), with the alternative notation $f^*(\{y\})$.

(x) Let $[X, Y, f\,]$ be a function. Generally, $f_*(X) = R_f \subseteq Y$. In this case, we say the map f is *from* X *into* Y. When $f_*(X) = Y$, we say the map f is *from* X *onto* Y or *surjective* or *surjection*. We call f *bijective* or *bijection* if f is both injective (one-to-one) and surjective (onto).

(xi) Let $f \subseteq X \times Y$ and $g \subseteq Y \times Z$ be two binary relations. Then the *composition* (or *composite*) *of f with g* is defined as

$$g \circ f = \{(x, z) \in X \times Z\colon\ \exists y\colon\ (x, y) \in f,\ (y, z) \in g\}.$$

The composition of f with g is most frequently used when $[X, Y, f]$ and $[Y, Z, g]$ are two functions and with $R_f \cap D_g \neq \emptyset$.

(xii) A consistent function $[X, Y, I]$, with $X \subseteq Y$, is called a *unity* or *identity function* if for all $x \in X$, $I(x) = x$. □

2.2 Example. For a fixed subset $A \subseteq X$, define the *indicator function* as a consistent map $[X, \mathbb{R}, \mathbf{1}_A]$ with

$$\mathbf{1}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \in A^c. \end{cases}$$

Then, $[X, \mathbb{R}, \mathbf{1}_A]$ is an into map, and the map $[X, \{0, 1\}, \mathbf{1}_A]$ is surjective. □

2.3 Definition. Let $f\colon\ X \to Y$ and let $A \subseteq X$. Then define

$$\mathrm{Res}_A f = \{(x, y) \in (A \times Y) \cap f\}.$$

This function is called the *restriction of f to A*. On the other hand, the function f is called an *extension of* the function $\mathrm{Res}_A f$.

In general, under a *consistent extension* of $[X, Y, f]$ we understand an expansion of X to a superset X' and some reasonable definition of f on $X' \backslash X$. We come across this problem in Chapter 5 regarding an extension of a set function. □

2.4 Example. Consider $[\mathbb{R}, [-1, 1], \sin]$ which is surjective (onto) but not injective (one-to-one). Take a restriction of function $[\mathbb{R}, [-1, 1], \sin]$ to one of the largest subsets A of $\mathbb{R}$ where $[\mathbb{R}, [-1, 1], \sin]$ is monotone increasing. It is plausible to set $A = [-\frac{\pi}{2}, \frac{\pi}{2}]$, because this set is also symmetric about the Y-axis. Then $[A, [-1, 1], \mathrm{Res}_A \sin]$ is obviously bijective and its inverse is the well-known function $[[-1, 1], A, \arcsin]$. □

2.5 Remark. Let $[X, Y, f]$ be a single-valued function such that for some $y \in R_f$, $f^*(\{y\}) = \{x_1, x_2, x_3\} \subseteq X$. Consider the composition $f_* \circ f^*$ and find that

$$f_* \circ f^*(\{y\}) = f_*(\{x_1, x_2, x_3\}) = \{y\}.$$

Thus, if f is single-valued, the restriction of $f \circ f^{-1}$ to R_f is the identity function with the domain $D_{f \circ f^{-1}} = R_f$. However, $f^{-1} \circ f$ need not be a single-valued function at all (show it). $f^{-1} \circ f$ is the identity function only when f is injective. □

PROBLEMS

2.1 Find the image of $[-3, 5)$ under $\mathbf{1}_{(1,2]}$.

2.2 Find the inverse image of $(\frac{1}{2}, 4]$ under $\mathbf{1}_{(1,2]}$.

2.3 Composition:

$a)$ Show that the composition operation $\circ$ is associative.
$b)$ Show that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.
$c)$ Show that $D_{g \circ f} = D_f \cap f^*(D_g)$.

2.4 Show the equivalence of the following statements:

$a)$ f is injective.

$b)$ $f_*(A \cap B) = f_*(A) \cap f_*(B)$.

$c)$ For every pair A and B, of disjoint sets, $f_*(A) \cap f_*(B) = \emptyset$.

What can go wrong with $b)$ if f is not injective?

In the following problems we assume that f is a map from $X = D_f$ into Y.

2.5 Show that $A \subseteq X \Rightarrow A \subseteq f^*(f_*(A))$.

2.6 Show that $\forall B \subseteq Y,\ \ f_*(f^*(B)) \subseteq B$.

2.7 Show that $[X, Y, f]$ is surjective if and only if

$$f_*(f^*(B)) = B,\ \forall B \subseteq Y.$$

2.8 Show that the map $[\mathcal{P}(R_f), \mathcal{P}(D_f), f^*]$ is injective.

2.9 Is the map $[\mathcal{P}(D_f), \mathcal{P}(R_f), f_*]$ injective?

2.10 Is the map $[\mathcal{P}(R_f), \mathcal{P}(D_f), f^*]$ surjective?

NEW TERMS:

Leibnitz, Gottfried von, 14
Bernoulli, Johann, 14
Euler, Leonard, 14
Fourier, Joseph, 14
Dirichlet, Gustav, 14
Cartesian (direct) product, 14
binary relation, 14
function, 14
single-valued function, 14
map [X, Y, f], 15
mapping, 15
transformation, 15
domain, 15
range, 15
codomain, 15
image of a point, 15
multivalued function, 15
branch of a function, 16
inconsistent function, 16
consistent function, 16
image of a set under a function, 16
function [$\mathcal{P}(X), \mathcal{P}(Y), f_*$], 16
inverse [$\mathcal{P}(R_f), \mathcal{P}(D_f), f^*$], 17
inverse image of function f_*, 17
pre-image, 17
inverse of function f, 17
function [$\mathcal{P}(\mathcal{P}(X)), \mathcal{P}(\mathcal{P}(Y)), f_{**}$] , 17
inverse [$\mathcal{P}(\mathcal{P}(R_f)), \mathcal{P}(\mathcal{P}(X)), f^{**}$], 17
one-to-one map (function), 17
injective map (function), 17
invertible map (function), 17
into map (function), 17
onto map (function), 17
surjective map (function), 17
bijective (onto and one-to-one) map (function), 17

composition of binary relations, 18
composition of maps, 18
unity function, 18
identity function, 18
indicator function, 18
restriction of a map (function), 18
extension of a map (function), 18
consistent extension of a map (function), 18

3. SET OPERATIONS UNDER MAPS

The most remarkable property of the inverse of a function is that it "preserves" all set operations (see Problem 3.6). The function itself, as we show, does not. The main theorems in this section are proved for special cases of surjective maps; the rest are left for the reader. Also, without loss of generality, consistency is assumed throughout.

3.1 Theorem. *Let* $[X, Y, f]$ *be a surjective map and let* $B \subseteq Y$. *Then*

$$[f^*(B)]^c = f^*(B^c)\,.$$

Proof. We prove an equivalent statement, $f^*(B) + f^*(B^c) = X$, that is, we show that

(i) $f^*(B)$ and $f^*(B^c)$ are disjoint and (ii) $f^*(B)$ complements $f^*(B^c)$ up to X.

We start with

(i) Suppose $f^*(B)$ and $f^*(B^c)$ have a common point x. Then there is $y_1 \in B$ such that $f(x) = y_1$ and $y_2 \in B^c$ such that $f(x) = y_2$. Thus, $y_1 \neq y_2$ and f is not a single-valued function. (See Figure 3.1.)

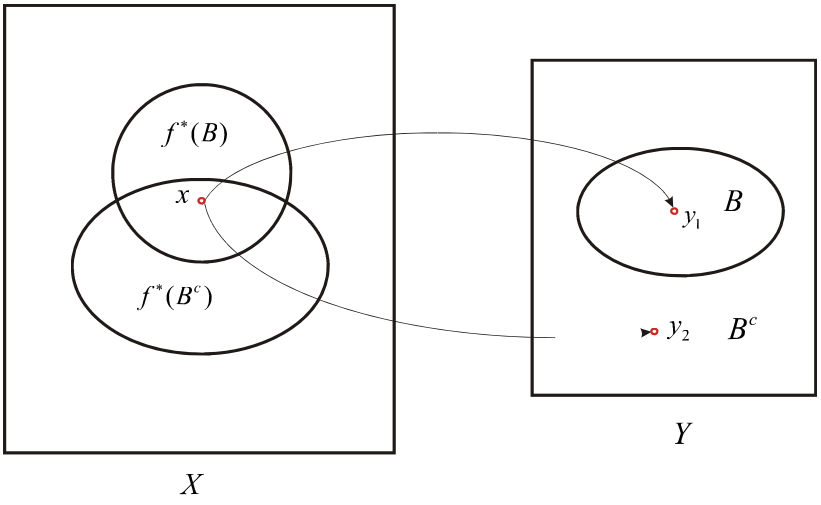


Figure 3.1

(ii) If $f^*(B)$ does not complement $f^*(B^c)$ up to X, there will be at least one point x that does not belong to either of these sets (for they are disjoint as shown above). This is an obvious contradiction, because it follows that $f(x) \notin Y$. (See Figure 3.2 below.) □

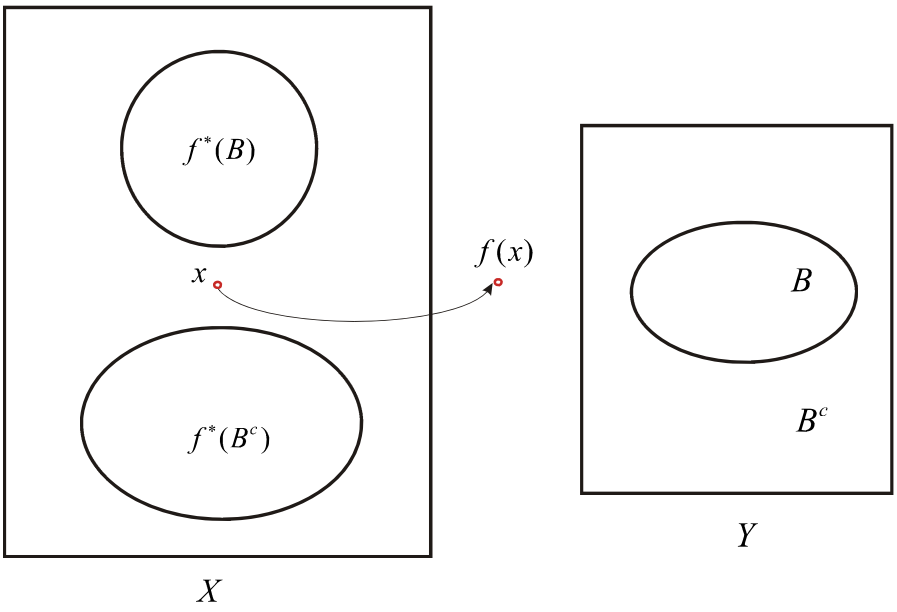


Figure 3.2

3.2 Example. Let $[X, Y, f]$ be a function. Then $[f^*(Y)]^c = X^c = \emptyset$. On the other hand, setting $B = Y$, by Problem 3.1, we obtain

$$f^*(Y^c) = f^*(\emptyset);$$

that is, $f^*(\emptyset) = \emptyset$. □

3.3 Theorem. *Let* $[X, Y, f]$ *be a surjective map. Then* $B_1 \subseteq B_2 \subseteq Y$ *implies that* $f^*(B_1) \subseteq f^*(B_2)$.

Proof. Suppose that $f^*(B_1)$ is not a subset of $f^*(B_2)$. This implies the existence of a point x that belongs to $f^*(B_1)$ and does not belong to $f^*(B_2)$. Therefore, there is exactly one point $y \in B_1$

with $f(x) = y$. On the other hand, because $x \notin f^*(B_2)$, $f(x)$ cannot belong to B_2. But it must, because $f(x) = y \in B_1 \subseteq B_2$. (See Figure 3.3 below.) Hence, our assumption above was wrong. □

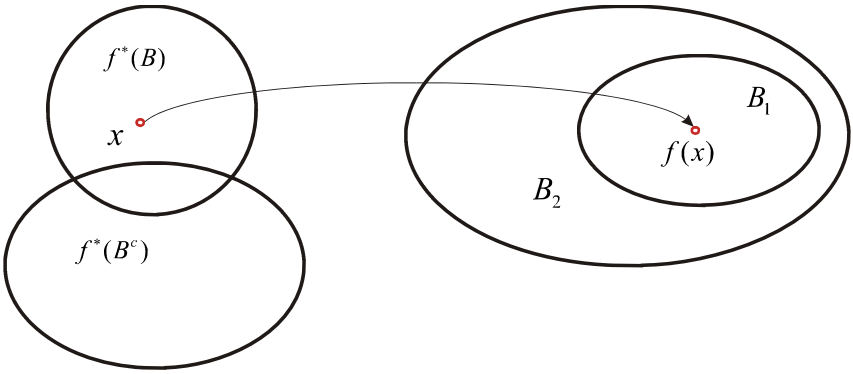


Figure 3.3

3.4 Theorem. *Let $f\colon X \to Y$ be an onto map and let $\{B_i : i \in I\}$ be an indexed family of subsets of Y. Then,*

$$f^*\left(\bigcup_{i\in I} B_i\right) = \bigcup_{i\in I} f^*(B_i).$$

Proof.

(i) We prove that $\bigcup_{i\in I} f^*(B_i) \subseteq f^*\left(\bigcup_{i\in I} B_i\right)$.

Let $x \in \bigcup_{i\in I} f^*(B_i)$. Then there is an index $i_0 \in I$ such that $x \in f^*(B_{i_0})$. Because $B_{i_0} \subseteq \bigcup_{i\in I} B_i$, by Theorem 3.3, $f^*(B_{i_0}) \subseteq f^*\left(\bigcup_{i\in I} B_i\right)$, which implies that $x \in f^*\left(\bigcup_{i\in I} B_i\right)$.

(ii) We show the validity of the inverse inclusion,

$$f^*\left(\bigcup_{i\in I} B_i\right) \subseteq \bigcup_{i\in I} f^*(B_i).$$

Let $x \in f^*\left(\bigcup_{i\in I} B_i\right)$. Then $f(x) \in \bigcup_{i\in I} B_i$. Therefore, there is an index $i_0 \in I$ such that $f(x) \in B_{i_0}$ if and only if $\{f(x)\} \subseteq B_{i_0}$. By Theorem 3.3, it follows that $f^*\{f(x)\} \subseteq f^*(B_{i_0})$. Because $x \in f^*(\{f(x)\})$, we have

$$\{x\} \subseteq f^*\{f(x)\} \subseteq f^*(B_{i_0}) \subseteq \bigcup_{i\in I} f^*(B_i). \qquad \square$$

PROBLEMS

3.1 Prove Theorem 3.1 under the condition that f is an into map. *Let* $[X, Y, f]$ *be an into map and let* $B \subseteq Y$. *Then*

$$[f^*(B)]^c = f^*(B^c).$$

3.2 Prove Theorem 3.3 under the condition that f is an into map: Let $[X, Y, f]$ be an into map. Then $B_1 \subseteq B_2 \subseteq Y$ implies that $f^*(B_1) \subseteq f^*(B_2)$.

3.3 Generalize Theorem 3.4 when f is an into map: Let $f\colon X \to Y$ be an into map and let $\{B_i : i \in I\}$ be an indexed family of subsets of Y. Then,

$$f^*\left(\bigcup_{i\in I} B_i\right) = \bigcup_{i\in I} f^*(B_i).$$

3.4 Let $[X, Y, f]$ be an into map and let $\{B_i : i \in I\}$ be an indexed family of subsets of Y.

a) Prove that $f^*\left(\bigcap_{i\in I} B_i\right) = \bigcap_{i\in I} f^*(B_i)$.

b) If B_1 and B_2 are disjoint, show that their inverse images under f are also disjoint.

c) If $\{B_i : i \in I\}$ is a pairwise disjoint family, show that

$$f^*(\sum_{i\in I} B_i) = \sum_{i\in I} f^*(B_i).$$

3.5 Show that $f^*(A\backslash B) = f^*(A)\backslash f^*(B)$.

3.6 The results above prove that all set operations are closed under the inverses of maps. Show that not all set operations are closed under maps as per the following.

a) Show that maps preserve inclusions.

b) Show that maps preserve unions.

c) Show that maps do not preserve intersections; specifically, show that

$$f_*\left(\bigcap_{i\in I} A_i\right) \subseteq \bigcap_{i\in I} f_*(A_i)$$

and that the inverse inclusion need not hold. Explain the latter with and without a counterexample.

d) Do maps preserve the difference?

3.7 Let $[X, Y, f]$ be a map and let $A \subseteq Y$. Show that

$$\mathbf{1}_A \circ f(x) = \mathbf{1}_{f^*(A)}(x).$$

3.8 Prove the following properties of the indicator function defined on a nonempty set Ω.

(i) $\mathbf{1}_{A\cap B} = \min\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A\mathbf{1}_B$.

(ii) $\mathbf{1}_{A\cup B} = \max\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A\cap B}$.

(iii) $\mathbf{1}_{A+B} = \mathbf{1}_A + \mathbf{1}_B$.

(iv) $\mathbf{1}_{\Sigma_{i\in I} A_i} = \sum_{i\in I} \mathbf{1}_{A_i}$.

(v) $\mathbf{1}_{A^c} = \mathbf{1} - \mathbf{1}_A$.

(vi) $A \subseteq B \Rightarrow \mathbf{1}_A \leq \mathbf{1}_B$.

(vii) $\mathbf{1}_{\bigcup_{i\in I} A_i} = \sup\{\mathbf{1}_{A_i}\colon i \in I\}$, $\mathbf{1}_{\bigcap_{i\in I} A_i} = \inf\{\mathbf{1}_{A_i}\colon i \in I\}$.

3.9 Let $\{A_n\}$ be a sequence of subsets of Ω. Show that the function $\overline{\lim} 1_{A_n}$ is the indicator function of the set $\overline{\lim}\, A_n$ (i.e., $\mathbf{1}_{\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m}$) and that the function $\underline{\lim}\, 1_{A_n}$ is the indicator function of the set $\underline{\lim}\, A_n$ (i.e., $\mathbf{1}_{\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m}$).

3.10 Prove that $\lim_{n\to\infty} A_n$ exists if and only if $\lim_{n\to\infty} \mathbf{1}_{A_n}$ exists.

3.11 Let $[X, X', F]$ be a bijective map and let τ and τ' be respective collections of subsets of X and X' such that

$$F^{**}(\tau') \subseteq \tau \quad \text{and} \quad F_{**}(\tau) \subseteq \tau'.$$

[See Definition 2.1(vii) and $(viii)$]. Show that

$$F^{**}(\tau') = \tau \text{ and } F_{**}(\tau) = \tau'.$$

3.12 Let $[X, Y, f]$ be a consistent map. Denote $F = f_* \circ f^*$. Then, the map $[\mathcal{P}(Y), \mathcal{P}(Y), F]$ is such that $F = I_*$ (I is the identity map) if and only if $[X, Y, f]$ is surjective.

3.13 Let $[X, Y, f]$ be a surjective map. Show that $[\mathcal{P}(X), \mathcal{P}(Y), f_*]$ is surjective.

4. RELATIONS AND WELL-ORDERING PRINCIPLE

In Definition 2.1(ii) we introduced the concept of a binary relation R as an arbitrary subset of $A \times B$. In the special case when $\mathrm{R} \subseteq A \times B$ and $A = B$, we call R a *binary relation on* A. We sometimes use as notation $a\mathrm{R}b$ instead of $(a, b) \in \mathrm{R}$. This notation makes sense, for instance, if R is stipulated by $<$ or $\leq$ on some set. In addition, we also say that a pair (X, R) is a binary relation, where in fact R is a binary relation on a set X (a *carrier*). Now we consider some special relations.

4.1 Definitions. Let R be a binary relation on X.

(i) R is called *reflexive* if $\forall a \in X$, $(a, a) \in \mathrm{R}$ $[a\mathrm{R}a]$.

(ii) R is called *symmetric* if $(a, b) \in \mathrm{R} \Rightarrow (b, a) \in \mathrm{R}$ $[a\mathrm{R}b \Rightarrow b\mathrm{R}a]$.

(iii) R is called *antisymmetric* if $(a, b), (b, a) \in \mathrm{R} \Rightarrow a = b$ $[a\mathrm{R}b \wedge b\mathrm{R}a \Rightarrow a = b]$.

(iv) R is called *transitive* if $(a, b), (b, c) \in \mathrm{R} \Rightarrow (a, c) \in \mathrm{R}$ $[a\mathrm{R}b \wedge b\mathrm{R}c \Rightarrow a\mathrm{R}c]$.

(v) R is called an *equivalence on* X (denoted by symbol $\approx$ or E) if it is *reflexive*, *symmetric*, and *transitive*.

(Observe that the equivalence E on X partitions X into mutually disjoint subsets, called *equivalence classes.* A *partition* of X is a family of disjoint subsets of X whose union is a *decomposition* of X. The elements of X "communicate" only within these classes. In summary, every equivalence relation generates mutually disjoint classes. The converse is also true: an arbitrary partition of the carrier X generates an equivalence relation.)

(vi) R is called a *partial order* (denoted by the symbol $\preceq$) if it is *reflexive*, *antisymmetric*, and *transitive.*

(vii) If $\preceq$ is a partial order, it is called *linear* or *total* if every two elements of X are related or *comparable*; *that is,* $\forall a, b \in X$ either $a \preceq b$ or $b \preceq a$.

$(viii)$ Let X be an arbitrary set and let E ($\approx$) be an equivalence relation on X. For $t \in X$ denote

$$[x]_{\mathrm{E}} \ (= [x]_{\approx}) = \{y \in X : y\mathrm{E}x\}$$

and call it an *equivalence class modulo* E ($\approx$). The set of all equivalence classes

$$X|\mathrm{E} = X|_{\approx} \text{ (or } \{[x]_{\mathrm{E}}\})$$

is said to be the *quotient* (or *factor*) *set of* X *modulo* E (or $\approx$). It is easily seen that a quotient set of X is also a partition of X.

Note that $x \mapsto [x]_{\mathrm{E}}$ is a function assigning to each $x \in X$, an equivalence class $[x]_{\mathrm{E}}$. We denote this function most often by π_{E} (or sometimes $\pi_{\approx}$) and call it the *projection of* X *on its quotient by* E (or $\approx$). In other words, for each $x \in X$, $\pi_{\mathrm{E}}(x) = [x]_{\mathrm{E}}$. □

4.2 Examples.

(i) $(\mathbb{R}, =)$ is an equivalence relation. Therefore, every real number as a singleton represents an equivalence class.

(ii) $(\mathbb{R}, \leq)$ is a linear order.

(iii) Congruent triangles on a plane offer an equivalence relation on the set of all triangles. [Two sets A and B are called *congruent* if there exists an "isometric" bijective map $f\colon A \to B$; i.e., f must preserve the distance for every pair of points $a, b \in A$ and their images $f(a), f(b) \in B$.]

(iv) $(\mathbb{R}^2, \leq)$ is not a linear order if we define "$\leq$" as $(a_1, b_1) \leq (a_2, b_2)$ if and only if $a_1 \leq a_2 \ \wedge \ b_1 \leq b_2$. To make this relation a linear order we can define, for instance, $(a_1, b_1) \leq (a_2, b_2)$ if and only if $\|(a_1, b_1)\| \leq \|(a_2, b_2)\|$, where $\|(a, b)\|$ is Euclidean distance of point (a, b) from the origin.

(v) Let $\mid$ be the relation on $\mathbb{N}$ such that $n \mid m$ if and only if n divides m (without a remainder). It can be shown that $(\mathbb{N}, \mid)$ is a partial order but not a linear order. (See Problem 4.5.)

(vi) Let p be a fixed integer greater than or equal to 2. Two integers a and b are called *congruent modulo* p if $a - b$ is divisible by p (without remainder); in notation we write $p \mid a - b$ or $a \equiv b \pmod{p}$. The number p is called the *modulus of congruence.* Define

$$[m]_p = \{n \in \mathbb{Z}\colon\ m \equiv n \pmod{p}\}\ (m \in \mathbb{Z}).$$

In other words,

$$[m]_p\ = \{n \in \mathbb{Z}\colon\ \exists k \in \mathbb{Z}\colon\ n = kp + m\}.$$

Then any two integers m and n are related in terms of $[\,\cdot\,]_p$ if and only if $n \in [m]_p$. This is an equivalence relation. (Show it; see Problem 4.1.)

(vii) Let X be a nonempty set and $\mathrm{R} \subseteq X \times X$ be a binary relation. Taking for R the diagonal $D = \{(x, x)\colon x \in X\}$ we have with (X, D) the "smallest" (by the contents of elements of $X \times X$) equivalence relation on X, where each element forms a singleton-class, and D partitions X into the "number" of classes corresponding to the quantity of all elements of X. The "largest" equivalence relation on X is obviously $\mathrm{R} = X \times X$ itself and it consists of the single class.

$(viii)$ Any consistent function $[X, Y, f\,]$ (i.e., a function with $X = D_f$) generates an equivalence relation on its domain X partitioning X into disjoint subsets. Define the binary relation E_f ($\approx_f$) on X as

$$x_1 \mathrm{E}_f x_2\ \Leftrightarrow\ f(x_1) = f(x_2).$$

Then, it is readily seen that E_f is an equivalence relation on X, referred to as the *equivalence kernel of function* f. Formally, for every point $y \in R_f = f_*(X)$, the preimage $f^*(\{y\})$ is an equivalence class in X and $X|\mathrm{E}_f = \{f^*(\{y\})\colon y \in R_f\}$ is the quotient set of X modulo E_f. Furthermore, $\sum_{y \in R_f} f^*(\{y\})$ is a decomposition of X.

For instance, the function $f(x) = x^2$ generates a partition of $\mathbb{R}$ into a collection of subsets of the form $\{-a, a\}$, for $a > 0$, along with $\{0\}$, which is a factor (quotient) set of $\mathbb{R}$ modulo E_{x^2}.

Another example is the function

$$\left[X = \mathbb{R}\backslash\left\{\tfrac{\pi(2n-1)}{2} : n \in \mathbb{Z}\right\}, \mathbb{R}, \tan\right].$$

Let

$$A_y = \tan^*(\{y\}) = \{\arctan y + \pi n : n \in \mathbb{Z}\} = [\arctan y]_{\mathrm{E}_{\tan}}.$$

Then, $\mathrm{E}_{\tan}$ is the equivalence kernel of the function tan and

$$X|\mathrm{E}_{\tan} = \{\tan^*(\{y\}) : y \in \mathbb{R}\}$$

is the quotient set of X modulo $\mathrm{E}_{\tan}$. Furthermore,

$$X = \textstyle\sum_{y\in\mathbb{R}} A_y. \qquad \square$$

The last discussion about an equivalence relation generated by a function yields some important results and notions we use in the forthcoming materials of Chapter 6.

We demonstrated in Example 4.2($viii$) that any function on X generates an equivalence relation. However, the following proposition states that the converse is also true, namely that any equivalence relation E is the equivalence kernel of some function, as per Proposition 4.3.

4.3 Proposition. *Let* E *be an equivalence relation on a nonempty set* X. *Then the projection map* $[X, X|\mathrm{E}, \pi_{\mathrm{E}}]$ (*defined as* $\pi_{\mathrm{E}}(x) = [x]_{\mathrm{E}}$) *is a surjective map with* E *being the equivalence kernel of* π_{E}. $\square$

Proof. From the definition of π_{E} it follows that π_{E} is surjective. To claim that E is the equivalence kernel of π_{E}, we need to show that $\pi_{\mathrm{E}}(x_1) = \pi_{\mathrm{E}}(x_2)$ if and only if $x_1 \mathrm{E} x_2$.

Let $\pi_{\mathrm{E}}(x_1) = \pi_{\mathrm{E}}(x_2)$. Because $x\mathrm{E}x$, $x \in [x]_{\mathrm{E}}$ and thus, by the assumption $\pi_{\mathrm{E}}(x_1) = \pi_{\mathrm{E}}(x_2)$, $x_1 \in [x_2[_{\mathrm{E}}$. This proves that $x_1 \mathrm{E} x_2$.

Now let $x_1 \mathrm{E} x_3$. If $x_2 \in [x_1]_\mathrm{E}$, then $x_2 \mathrm{E} x_1$ and thus, by transitivity, $x_2 \mathrm{E} x_3$; that is, $x_2 \in [x_3]_\mathrm{E}$. Therefore, $[x_1]_\mathrm{E} \subseteq [x_3]_\mathrm{E}$. The inverse inclusion, and thus the equality, is due to the symmetry of E. Hence, $\pi_\mathrm{E}(x_1) = \pi_\mathrm{E}(x_3)$. □

Proposition 4.3 asserts that the projection π_E is an example of a surjective function defined on X and with the range $X|\mathrm{E}$.

Now suppose E is an equivalence relation on a set X and $[X, Y, f]$ is any function whose equivalence kernel is E. The following theorem claims that there is a unique "intercessor" $\boldsymbol{f}$ between the quotient set $X|\mathrm{E}$ and the codomain Y of f.

4.4 Theorem. *Let* $[X, Y, f]$ *be a consistent function with* E_f *being its equivalence kernel. Then, there is a unique function* $[X|\mathrm{E}_f, Y, \boldsymbol{f}]$ *such that* $f = \boldsymbol{f} \circ \pi_{\mathrm{E}_f}$. □

The reader shall be able to take care of this theorem (Problem 4.10) as well as of Corollary 4.5 (Problems 4.11 and 4.12).

We call $\boldsymbol{f}$ the *reducer of* f.

4.5 Corollary. *Let* $[X, Y, f]$ *be a function and let* E_f *denote its equivalence kernel. Then* : (a) *there is a unique reducer function* $[X|\mathrm{E}_f, Y, \boldsymbol{f}]$ *of* f *such that* $\boldsymbol{f}$ *is injective and* f *can be represented as a composition* $f = \boldsymbol{f} \circ \pi_{\mathrm{E}_f}$ of $[X, X|\mathrm{E}_f, \pi_{\mathrm{E}_f}]$ *and* $[X|\mathrm{E}_f, Y, \boldsymbol{f}]$; (b) *if* f *is surjective*, *the reducer* $\boldsymbol{f}$ *is bijective.* □

4.6 Example. To illustrate Theorem 4.4 and Corollary 4.5 consider $[\mathbb{R}, [-1, 1], \cos]$. If $\mathrm{E}_{\cos}$ denotes the equivalence kernel of the cosine function, then

$$\mathbb{R}|\mathrm{E}_{\cos} = \{\cos^*\{y\} : y \in [-1, 1]\}$$

is the quotient set of $X = \mathbb{R}$ modulo $\mathrm{E}_{\cos}$ such that

$$\cos^*\{y\} = \{\pm \arccos y + 2\pi n : n \in \mathbb{Z}\}$$

and

$$\pi_{\mathrm{E}_{\cos}}(x) = [x]_{\mathrm{E}_{\cos}} = \{\pm x + 2\pi n : n \in \mathbb{Z}\}$$

$$= \cos^*\{\cos x\}$$

making $[\mathbb{R}, \mathbb{R}|\mathrm{E}_{\cos}, \pi_{\mathrm{E}_{\cos}}]$ the projection of $X = \mathbb{R}$ on its quotient by $\mathrm{E}_{\cos}$. The reducer $\boldsymbol{f}$ of f is denoted by $[\mathbb{R}|\mathrm{E}_{\cos}, [-1, 1], \mathbf{cos}]$ and by Corollary 4.5, it is bijective. Also, by Corollary 4.5,

$$\mathbf{cos} \circ \pi_{\mathrm{E}_{\cos}}(x) = \mathbf{cos}(\pm x + 2\pi n : n \in \mathbb{Z}) = \cos x. \qquad \square$$

Now we turn to a discussion on the partial order relation and all relevant notions and theorems, which we apply throughout the rest of the book.

4.7 Definitions. Let $(A, \preceq)$ be a partial order and let $B \subseteq A$. Clearly, $(B, \preceq)$ is also a partial order.

(i) The partial order $(B, \preceq)$ is called a *chain in* $(A, \preceq)$ if it is linear.

(ii) An element $b_0 \in B$ is called a *minimal element* of B (relative to $\preceq$) if for each $b \in B$ with $b \preceq b_0$, $b = b_0$ (compared with the *smallest element* b_0, which is $\preceq b$ for all $b \in B$).

(iii) An element $b_\infty \in B$ is called a *maximal element* of B (relative to $\preceq$), if for each $b \in B$, with $b_\infty \preceq b$, it holds true that $b = b_\infty$ (compared with the *largest element* b_∞, which is such that $b \preceq b_\infty \ \forall b \in B$).

(Observe that the difference between a minimal element and the smallest element of a set is as follows. A minimal element b_0 is $\preceq b \in B$ whenever b_0 is comparable with some b. In addition, the smallest element is comparable with all elements of B.)

(iv) An element $u \in A$ is said to be an *upper bound of* B if $b \preceq u \ \forall b \in B$. An element $l \in A$ is said to be a *lower bound of* B if $l \preceq b \ \forall b \in B$. If B has lower and upper bounds then B is called *bounded* (or $\preceq$ *-bounded*).

(v) If the set of upper bounds of B has a smallest element u_0 then this element is called the *least upper bound of set* B [abbreviated lub(B)] or *supremum* [sup(B)]. Similarly, if the set of all lower bounds has a largest element l_∞ then it is called the *greatest lower bound of the set* B [in notation glb(B)] or *infimum* [inf(B)].

[For instance, 0 is the glb$((0,1))$ or inf$(0,1)$ in $(\mathbb{R},\ \leq)$, whereas an lub of the set $[1, \sqrt{2}] \cap \mathbb{Q}$ does not exist in $(\mathbb{Q},\ \leq)$.]

(vi) Let B contain at least two points. The partial order $(B,\ \preceq)$ is called a *lattice* if the supremum and infimum of every two-element subset of B are also elements of B. In notation: if $A = \{x, y\},$ then

$$x \vee y :\ = \sup\{x, y\} = \sup(A) \in B$$

and

$$x \wedge y :\ = \inf\{x, y\} = \inf(A) \in B.$$

The symbols $\vee$ and $\wedge$ are referred to as *join* and *meet* lattice operations, respectively. □

4.8 Examples.

(i) Let $B = \{1, 3, 3^2, \dots, 3^n, \dots\}$. Then $(B,\ |\ )$ [where $|$ is the relation in Example 4.2(v)] is a chain in $(\mathbb{N},\ |\ )$.

(ii) Let $B = \{2, 3, 4, \dots\}$ and consider the relation $|$ on B. In terms of this relation, the set of all prime numbers $\{2, 3, 5, 7, 11, \dots\}$ is the set of all minimal elements, and there is no smallest element in $B,$ because there is no minimal element related to all other elements. B does not have a maximal element either.

(iii) Consider the partial order $(\mathcal{P}(\Omega),\ \subseteq)$. It is obvious that for an arbitrary indexed family $\mathcal{A} = \{A_i \subseteq \Omega : i \in I\} \subseteq \mathcal{P}(\Omega),$ it is true that $\sup \mathcal{A} = \bigcup_{i \in I} A_i \in \mathcal{P}(\Omega)$ and $\inf \mathcal{A} = \bigcap_{i \in I} A_i \in \mathcal{P}(\Omega)$. In particular, it holds true for pairs of subsets. Thus, $(\mathcal{P}(\Omega),\ \subseteq)$ is a lattice. □

4.9 Definition. A linear order $(A,\ \preceq)$ is said to be *well-ordered* if every nonempty subset of A has a smallest element in the sense of the same order $\preceq$. □

4.10 Example. Let $\mathbb{R}$ be the set of all real numbers and consider the relation $(\mathbb{R},\ \leq)$ which is clearly a linear order. However, $\mathbb{R}$ is not well-ordered by $\leq,$ for there are nonempty subsets containing no smallest element, such as $(0, 1)$. But $(\mathbb{N},\ \leq)$ is well-ordered. □

Can all sets be well-ordered? This is one of the fundamental questions in set theory posed by Georg Cantor in the 1870's. Cantor considered it obvious that every set can indeed be well-ordered. At that time set theory was not well-established yet. In 1908, Ernst Zermelo formulated his *axiom of choice* and showed in his paper, "Untersuchungen über die Grundlagen der Mengenlehre" (see [Zer]) that the axiom of choice is equivalent to the "well-ordering principle." The axiom of choice was included in an axiom scheme for set theory that was later (1922) strengthened by Adolf Abraham Fränkel in his paper, "Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre" (see [Frän]).

Zermelo and Fränkel introduced the following notions. Let $\mathcal{S}$ be a collection of sets. A function c defined on $\mathcal{S}$ is called a *choice function*, if for each $X \in \mathcal{S},\ c(X) \in X$. In other words, c assigns to each set exactly one element of the set. Or less formally, we can choose exactly one element from each set. Observe that if $\mathcal{S}$ is an indexed set, that is, $\mathcal{S} = \{X_i : i \in I\}$, then we have $f(i) = c(X_i) \in X_i$. The axiom of choice is formulated in this way:

Every system of sets has a choice function.

Zermelo proved that a *nonempty set A can be well-ordered if and only if its power set $\mathcal{P}(A)$ has a choice function.* (There will be a short discussion of the axiom of choice in the upcoming sections.)

4.11 Theorem (Zermelo). *The axiom of choice is equivalent to the well-ordering principle.* □

4.12 Examples.

(i) To illustrate a use of the axiom of choice, consider the following example. Let $[X, Y, f]$ be an onto map. We show that there exists a subset $A \subseteq X$ such that $\mathrm{Res}_A f\colon\ A \to Y$ is bijective. Let c be a choice function for the factor set $\{f^*(\{y\})\colon\ y \in Y\}$ of X modulo E_f. Then the set $A = \{c(f^*(\{y\}))\colon\ y \in Y\}$ has the desired property. In other words, we choose one x from $f^*(\{y\})$ for each y and the collection of all these x's is A.

(ii) Let $A = \{c(\tan^*\{y\}) = \arctan y : y \in \mathbb{R}\}$. Then, set A is $A = (-\frac{\pi}{2}, \frac{\pi}{2})$ and hence $[A, \mathbb{R}, \mathrm{Res}_A\tan]$ is a function such that it is injective and $(\mathrm{Res}_A\tan)^{-1} = \arctan$. □

One of the central results in set theory is Zorn's lemma [1935] belonging to a German-born American mathematician, Max August Zorn (1906 – 1993), which is widely used in set theory and which is also equivalent to the axiom of choice.

4.13 Lemma (Zorn). *If each chain in a partially ordered set A has an upper bound, then A has a maximal element.*

Another very useful result is Tukey's lemma (named so after the American mathematician and statistician John Tukey), which follows from Zorn's lemma. Tukey's lemma is sometimes referred to as the Teichmüller-Tukey lemma.

4.14 Lemma (Tukey). *Let Ω be a nonempty set and $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ be a family of subsets of Ω such that $A \in \mathcal{F}$ if and only if each finite subset of A is in $\mathcal{F}$. Then, $\mathcal{F}$ has a maximal element A_∞ in the sense of $\subseteq$, that is, there is no $B \in \mathcal{F}$ such that $A_\infty \subseteq B$.*

PROBLEMS

4.1 Show that the relation in Example 4.2 (vi) is an equivalence relation on $\mathbb{Z}$. Give the equivalence classes for $p = 4$.

4.2 Classify the following binary relations.
$a)$ Let Ω be a nonempty set. Define the relation $(\mathcal{P}(\Omega), \subseteq)$.
$b)$ Let $\Omega = \mathbb{R}^2\backslash(x, 0)$. Define R: $(a, b)\mathrm{R}(c, d) \Leftrightarrow ad = bc$.

4.3 The following theorem is a statement of the *principle of mathematical induction.*

Let $S(n)$ be a statement which is true or false, for $n = 1, 2, \ldots$. Let $S(1)$ be true and suppose whenever $S(n)$ is true implies that $S(n + 1)$ is true, $n = 1, 2, \ldots$. Then $S(n)$ is true for all n.

Prove it. (*Hint*: Use the well-ordering principle.)

4.4 Prove that $\sum_{i=1}^{n} i^2 = \frac{1}{6}n(n+1)(2n+1)$.

4.5 Show that $(\mathbb{N}, \mid)$ in Example 4.2(v) (defined as: $n \mid m$ if and only if n divides m) is a partial order relation.

4.6 Let l^0 denote the space of all summable real-valued sequences. In other words, a sequence $[\mathbb{N}, \mathbb{R}, f]$, in the form of $\{a_n\}$, is *summable*, if the series $\sum_{n=1}^{\infty} a_n$ converges. We define the partial order $\preceq$ on l^0 as $f \preceq g$ if and only if $f(n) \leq g(n)$ for all $n \in \mathbb{N}$. Is $(l^0, \preceq)$ a lattice?

4.7 Let $\mathcal{M}^0_{(n,n)}$ be the set of all non singular $n \times n$ matrices. We define the partial order $\preceq$ on $\mathcal{M}^0_{(n,n)}$ as $A = (a_{ij}) \preceq B = (b_{ij})$ if and only if and only if $a_{ij} \leq b_{ij}$ for all i and j. Is $(\mathcal{M}^0_{(n,n)}, \preceq)$ a lattice?

4.8 Is the set of all continuous real-valued functions a lattice?

4.9 Is the set of all real-valued polynomials a lattice?

4.10 Prove Theorem 4.4: *Let* $[X, Y, f]$ *be a consistent function with* E_f *being its equivalence kernel. Then there is a unique function* $[X|\mathrm{E}_f, Y, \boldsymbol{f}]$ *such that* $f = \boldsymbol{f} \circ \pi_{\mathrm{E}_f}$.

4.11 Prove part (a) of Corollary 4.5: *Let* $[X, Y, f]$ *be a function and let* E_f *denote its equivalence kernel. Then there is a unique reducer function* $[X|\mathrm{E}_f, Y, \boldsymbol{f}]$ *of* f *such that* $\boldsymbol{f}$ *is injective and* f *can be represented as a composition* $f = \boldsymbol{f} \circ \pi_{\mathrm{E}_f}$ of $[X, X|\mathrm{E}_f, \pi_{\mathrm{E}_f}]$ *and* $[X|\mathrm{E}_f, Y, \boldsymbol{f}]$.

4.12 Prove part (b) of Corollary 4.5: *Let* $[X, Y, f]$ *be a function and let* E_f *denote its equivalence kernel. Then if* f *is surjective, the reducer* $\boldsymbol{f}$ *is bijective.*

4.13 For the function $[\mathbb{R}, [-1, 1], \sin]$ render a similar analysis as in Example 4.6: give the equivalence kernel, the quotient set, the projection of $\mathbb{R}$ on its quotient set, and the reducer.

4.14 For the function $[[-\pi, \pi], [-1, 1], \sin]$ render a similar analysis as in Example 4.6: give the equivalence kernel, the quotient set, the projection of $[-\pi, \pi]$ on its quotient set, and the reducer. Is the reducer bijective?

4.15 For the function $[\mathbb{R}, [-1, \infty), f(x) = x^2 - 1]$ render a similar analysis as in Example 4.6: give the equivalence kernel, the quotient set, the projection of $\mathbb{R}$ on its quotient set, and the reducer.

4.16 Let $\left[\mathcal{C}_{\mathbb{R}}^{(1)}, \mathcal{C}_{\mathbb{R}}, \mathcal{D}^1\right]$ denote the differential operator [$\mathcal{D}^1 f(z)$ is defined as $\frac{d}{dz} f(z)$] on the set of all real-valued continuously differentiable functions. Give the equivalence kernel of $\mathcal{D}^1$, the quotient set, the projection of $\mathcal{C}_{\mathbb{R}}^{(1)}$ on its quotient set, and the reducer. Is the reducer bijective?

NEW TERMS:

binary relation on a set, 29
- reflexive, 29
- symmetric, 29
- antisymmetric, 29
- transitive, 29
- equivalence, 29
- partial order, 29
- linear (total) order, 29

comparable elements, 29
$[x]_{\mathrm{E}}$, equivalence class modulo E, 30
$X|\mathrm{E}$, quotient (factor) set, 30
factor (quotient) set, 30
π_{E}, projection of a set on its quotient, 30
congruence, 31
congruence modulo p, 31
modulus of congruence, 31
equivalent classes generated by a function, 31
equivalence kernel of a function, 31
$[X|\mathrm{E}_f, Y, \boldsymbol{f}]$, reducer, 33
chain, 34
minimal element, 34
smallest element, 34
maximal element, 34
largest element, 34
upper bound, 34
lower bound, 34
bounded set, 34
least upper bound (supremum), 34
supremum, 34
greatest lower bound (infimum), 34
infimum, 34
lattice, 35
join operation, 35
meet operation, 35
well-ordered set, 35

Cantor, Georg,, 36
Zermelo, Ernst, 36
axiom of choice, 36
well-ordering principle, 36
Fränkel, Adolf Abraham, 36
choice function, 36
Zermelo's theorem, 36
Zorn, Max, 37
Zorn's lemma, 37
Tukey, John, 37
Tukey's lemma, 37
principle of mathematical induction, 37

5. CARTESIAN PRODUCT

The idea of the Cartesian product (or, equivalently, *direct product*) primarily belongs to René Descartes who introduced this notion for two sets X and Y as a set of all ordered pairs $\{(x, y)\colon x \in X$ and $y \in Y\}$. Descartes was also the one who introduced the widely used Cartesian coordinate system related to the Cartesian product.

In Definition 2.1, we introduced the notion of the Cartesian product of finitely many sets. We extend this definition to arbitrarily many sets. We begin with sequences of sets.

5.1 Definitions.

(i) Let $\{Y_i\colon\ i = 1, 2, \dots\}$ be a sequence of arbitrary sets. Then the *Cartesian product of this sequence* is the set of all sequences

$$\prod_{n=1}^{\infty} Y_n = \{(a_1, a_2, \dots)\colon a_i \in Y_i,\ i = 1, 2, \dots\}$$

of elements from $Y_1,\ Y_2, \dots,$ respectively.

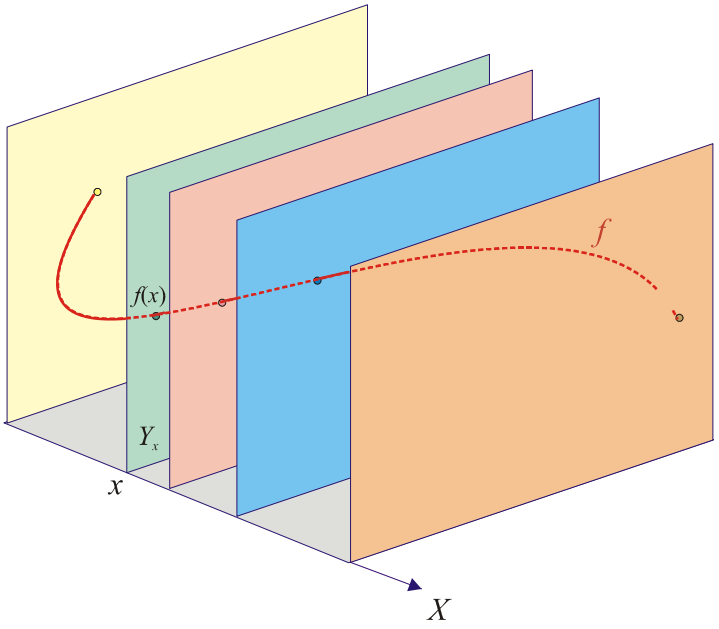


Figure 5.1

(ii) In the general case, let $\{Y_x : x \in X\}$ be an indexed family of sets.

Then the *Cartesian product* (see Figure 5.1 above) defined as

$$\prod_{x\in X} Y_x = \{f : X \to \bigcup_{x\in X} Y_x : f(x) \in Y_x\,,\ x \in X\}$$

is the collection of all functions defined on the index set X and valued in $\bigcup_{x\in X} Y_x$. Each such function is a choice function for the family $\{Y_x : x \in X\}$. For an $x \in X$, the set Y_x is called the xth *factor space*. □

5.2 Remarks.

(i) One of the basic questions that arises is this: *when is the Cartesian product nonempty?* Obviously, if at least one factor space $Y_k = \emptyset$, then $\prod_{x\in X} Y_x = \emptyset$. But if all $Y_x \neq \emptyset$, is the Cartesian product necessarily nonempty? Although the answer may seem obvious, we must turn to the axiom of choice. In other words, the Cartesian product of a family of sets is nonempty if and only if there exists at least one choice function for this family, which is guaranteed by the axiom of choice.

(ii) We said that the Cartesian product of the family of sets $\{Y_x : x \in X\}$ is the collection of all functions from X to $\bigcup_{x\in X} Y_x$. In particular, if $Y_x = Y$, for all $x \in X$, then the Cartesian product is the collection of all functions naturally denoted by Y^X [or by $\mathcal{F}(X; Y)$].

5.3 Example. As an illustration of a use of the Cartesian product we show that given a nonempty set X, the power set $\mathcal{P}(X)$ is "equipotent" with the set $\{0, 1\}^X$. So, let X be an arbitrary set. Then every subset $A \subseteq X$ can be associated with its indicator function $\mathbf{1}_A$. Conversely,

$$A = \{x \in X\colon \mathbf{1}_A(x) = 1\}.$$

Therefore, we can set a one-to-one correspondence between $\mathcal{P}(X)$ and the set of all indicator functions indexed with all subsets of X. On the other hand, the set of all such indicator functions is in fact the set of all (binary) functions of type β: $X \to \{0,1\}$. [Indeed, if β is a binary function, $\{x \in X\colon \beta(x) = 1\} = B$ is a subset of X. Thus, $\beta = \mathbf{1}_B$.] This set, by the above definition, is the Cartesian product of the family of sets $Y_x = \{0,1\}$, where the index x runs X, in notation, $\{0,1\}^X$. So, we have shown that $\mathcal{P}(X)$ is "equipotent" (i.e., in a bijective correspondence) with the set $\{0,1\}^X$ of all binary functions β: $X \to \{0,1\}$. □

5.4 Definitions (Projection Maps).

(i) Let $\{Y_x : x \in X\}$ be a collection of sets. Given an $\alpha \in X$, the map

$$\left[\prod_{x\in X} Y_x\,, Y_\alpha\,, \pi_\alpha\right]$$

is called the αth *projection map* if

$$\pi_\alpha(f) = f(\alpha),$$

where

$$f \in \prod_{x\in X} Y_x \text{ and } f(\alpha) \in Y_\alpha.$$

The point $f(\alpha)$ is called the αth *coordinate of* f and we recall that Y_α is the αth *factor space*. (See Figure 5.2 below.) Observe that $\pi^*_\alpha(\{f(\alpha)\}) \neq \{f\}$ but it contains $\{f\}$.

For instance, if $X = \{1,\dots,n\}$,

$$\pi^*_1(\{f(1)\}) = \{f(1)\} \times Y_2 \times \cdots \times Y_n\,.$$

In general,

$$\pi^*_\alpha(\{f(\alpha)\}) = \prod_{x\in X} \hat{Y}_x\,, \tag{5.4}$$

where

$$\hat{Y}_x = Y_x, \quad \text{for } x \neq \alpha, \text{ and } \hat{Y}_\alpha = \{f(\alpha)\}. \tag{5.4a}$$

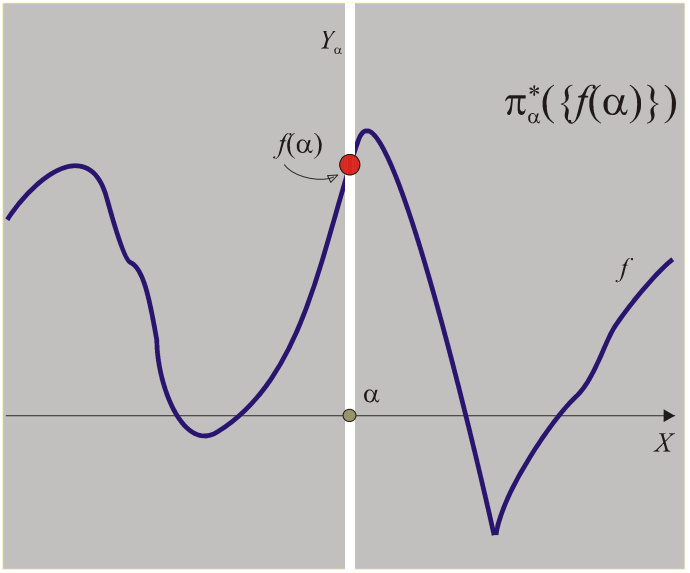


Figure 5.2

In a more familiar scenario of the 2-D Cartesian coordinate system the two figures below readily speak for themselves.

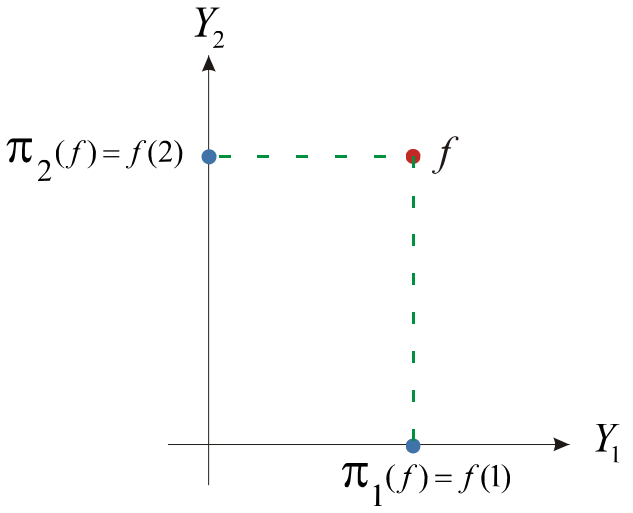


Figure 5.3

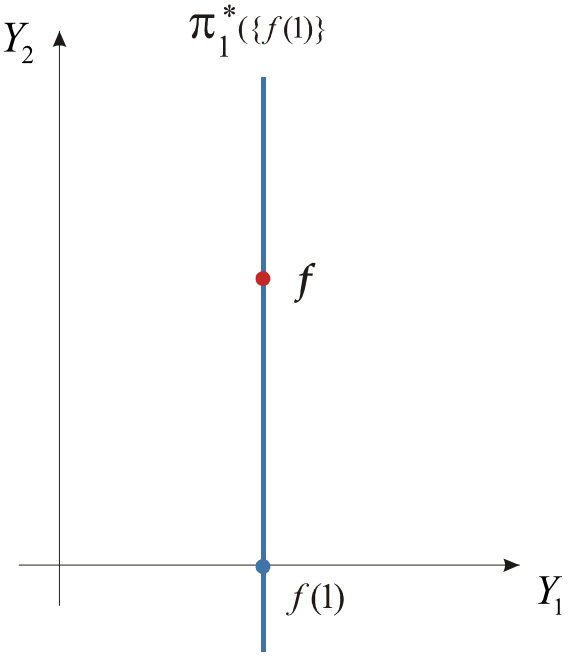


Figure 5.4

(ii) Obviously, from (5.4 - 5.4a) we have $\pi_\alpha^*(Y_\alpha) = \prod_{x\in X} Y_x$. In particular, in the latter two-dimensional case, $\pi_1^*(Y_1) = Y_1 \times Y_2$.

Now, if $B_\alpha \subseteq Y_\alpha$, then

$$\pi_\alpha^*(B_\alpha) = \prod_{x\in X} \hat{Y}_x\,,$$

where

$$\hat{Y}_x = Y_x, x \neq \alpha, \quad \text{and} \quad \hat{Y}_\alpha = B_\alpha.$$

The subset $\pi_\alpha^*(B_\alpha)$ of $\prod_{x\in X} Y_x$ is called a *unit cylinder with base* B_α. (See Figure 5.5 below.)

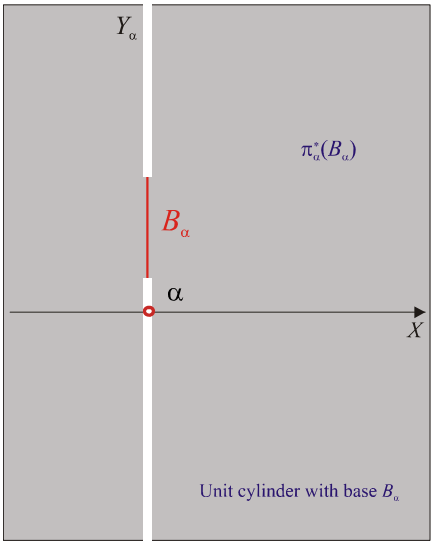


Figure 5.5

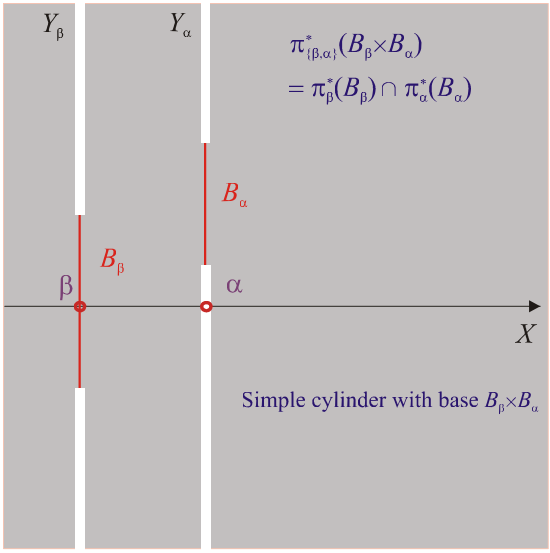


Figure 5.6

Notice that we should take pictorial illustrations like those in Figures 5.5 and 5.6 with extreme caution. For instance, if we take two disjoint sets $A_x, B_x \subseteq Y_x$, then because inverse images of disjoint sets are disjoint, so must be $\pi_x^*(A_x)$ and $\pi_x^*(B_x)$. However, their intersection will not look empty, for they will appear intersecting on the entire plane except on Y_x. This is because all Y_y's are nonoverlapping (which we need to remember whenever we come to draw a figure) and as depicted, they seem to form a solid area. We thus admit that these figures are useful, but imperfect, illustrations.

In the conventional, two-dimensional Cartesian system, a unit cylinder $\pi_1^*(B_1)$ with base B_1 looks like shown in Figure 5.7.

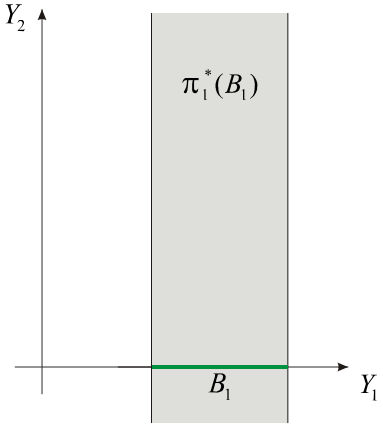


Figure 5.7

(iii) Now, for $\alpha, \beta \in X$, let $B_\alpha \subseteq Y_\alpha$ and $B_\beta \subseteq Y_\beta$. Then, obviously,

$$\pi_\alpha^*(B_\alpha) \cap \pi_\beta^*(B_\beta) = \prod_{x \in X} \hat{Y}_x \,,$$

where $\hat{Y}_x = Y_x, x \neq \alpha, \beta$, and $\hat{Y}_\alpha = B_\alpha, \hat{Y}_\beta = B_\beta$.

The set $\pi_\alpha^*(B_\alpha) \cap \pi_\beta^*(B_\beta)$ is called a *simple cylinder with base* $B_\alpha \times B_\beta$. (See Figure 5.6.) This construction of a simple cylinder can be extended to any finitely many sets $B_{\alpha_1}, \ldots, B_{\alpha_n}$.

(*iv*) Let $B_x \subseteq Y_x$, $x \in X$. The set

$$\prod_{x\in X} B_x = \bigcap_{x\in X} \pi_x^*(B_x)$$

is called a *rectangle* or *parallelepiped.*

If $A \subseteq X$ and

$$\hat{Y}_x = \begin{cases} B_x \subseteq Y_x, & x \in A \\ Y_x, & x \in X\backslash A, \end{cases}$$

then the set

$$\prod_{x\in X} \hat{Y}_x = \bigcap_{x\in A} \pi_x^*(B_x) \tag{5.4b}$$

is a (*rectangular*) *cylinder with base* $\prod_{x\in A} B_x$. A cylinder in (5.4b) is simple if set A is finite.

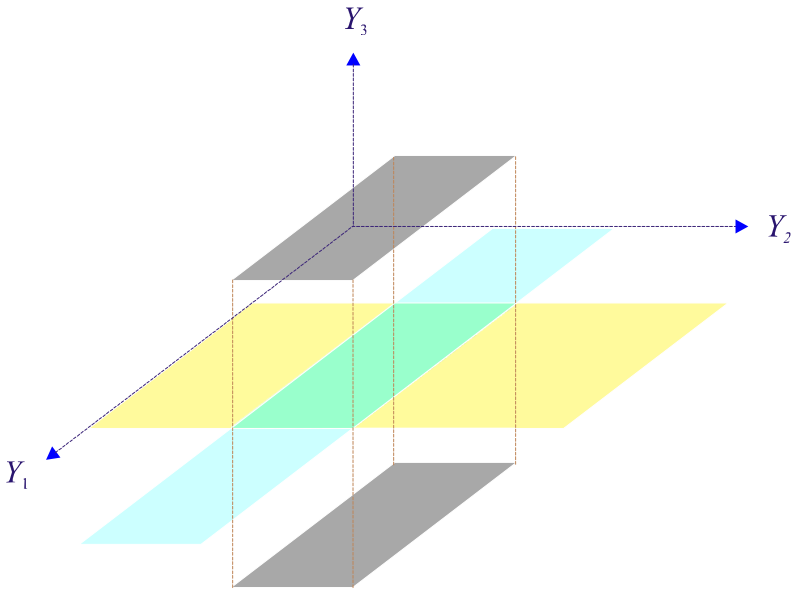


Figure 5.8

In a conventional Cartesian representation of the Euclidean three-dimensional space, common in calculus-based geometry, with $X = \{1, 2, 3\}$, the factor spaces Y_1, Y_2, Y_3 are drawn as mutually orthogonal axes. If we take $A = \{1, 2\}$ and B_1 and B_2 as intervals from Y_1 and Y_2, the cylinder with base $B_1 \times B_2$ will look like an infinite parallelepiped (with $B_3 = Y_3$). In Figure 5.8 above, the perpendicular strips are just "traces" of $\pi_1^*(B_1)$ and $\pi_2^*(B_2)$ on plane $Y_1 \times Y_2$.

(v) Now we introduce a more general notion of a projection map. Let $\{Y_x : x \in X\}$ be an arbitrary indexed family of sets and let $A \subseteq X$. Define

$$\left[\prod_{x\in X} Y_x\,,\ \prod_{\alpha\in A} Y_\alpha\,,\ \pi_A\right]$$

and call it the *A-projection map* if $\pi_A(f) = f_*(A)$. Specifically, if $A = \{\alpha\}$ we have $\pi_{\{\alpha\}}(f) = f_*(\{\alpha\})$ which, in contrast with definition (i), is a set, namely, a singleton.

In light of Definition (iv), the A-projection map π_A can be used in representation (5.4b) of a *rectangular cylinder with base* $\prod_{x\in A} B_x$ as $\pi_A^*\left(\prod_{x\in A} B_x\right)$, where B_x is a subset Y_x, for $x \in A$. Recall that a rectangular cylinder is *simple* if A is a finite subset of X. In particular, the rectangular simple cylinder depicted in Figure 5.7 (with base $B_1 \times B_2$) can be expressed as $\pi_{\{1,2\}}^*(B_1 \times B_2)$. If $A = \{x_1, \ldots, x_n\}$, a rectangular simple cylinder can also be expressed as

$$\pi_{\{x_1,\ldots,x_n\}}^*\left(\prod_{i=1}^{n} B_{x_i}\right) = \pi_{x_1}^*(B_{x_1}) \cap \cdots \cap \pi_{x_n}^*(B_{x_n}). \tag{5.4c}$$

(vi) With this more flexible notion of the A-projection map we can further generalize the notion of a cylinder. Let $\mathcal{A} \subseteq \prod_{\alpha\in A} Y_\alpha$. Then call $\pi_A^*(\mathcal{A})$ an *A-cylinder with base $\mathcal{A}$*. An A-cylinder is then *rectangular* if $\mathcal{A}$ is a rectangle with respect to $\prod_{\alpha\in A} Y_\alpha$, that is, $\mathcal{A} = \prod_{\alpha\in A} B_\alpha$ for $B_\alpha \subseteq Y_\alpha$.

In particular, if A is a finite subset of X then the rectangular cylinder becomes *simple* as in Figure 5.6 above. A simple A-cylinder is a *unit cylinder* if A is a singleton. (See Figure 5.5.) □

5.5 Examples.

(i) Let $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. Then,

$$\pi_{\{\alpha_n,\ldots,\alpha_n\}}(f) = f_*(\{\alpha_1, \ldots, \alpha_n\}) = \{f(\alpha_1), \ldots, f(\alpha_n)\},$$

and hence,

$$\pi^*_{\{\alpha_n,\ldots,\alpha_n\}}(\{f(\alpha_1), \ldots, f(\alpha_n)\}) = \bigcap_{i=1}^{n} \pi^*_{\alpha_i}(f(\alpha_i))$$

is an $\{\alpha_1, \ldots, \alpha_n\}$-simple cylinder with base $\{f(\alpha_1), \ldots, f(\alpha_n)\}$.

(ii) This example will be revisited in Section 5, Chapter 3. Consider as $\mathbb{R}^{\mathbb{R}}$ the collection of all real-valued functions on $\mathbb{R}$ regarded as the Cartesian product of $\mathbb{R}$'s. We select a "neighborhood" U_f of a point $f \in \mathbb{R}^{\mathbb{R}}$. First of all, a simple cylinder with base $(y_1 - \varepsilon_1, y_1 + \varepsilon_1) \times \cdots \times (y_k - \varepsilon_k, y_k + \varepsilon_k)$ obviously has the form

$$\pi^*_{\alpha_1}\left[(y_1 - \varepsilon_1, y_1 + \varepsilon_1)\right] \cap \cdots \cap \pi^*_{\alpha_k}\left[(y_k - \varepsilon_k, y_k + \varepsilon_k)\right], \quad (5.5)$$

where y_x is a point in $Y_x = \mathbb{R}$. This cylinder (in Figure 5.9 shaded with gray color including the red strips) becomes a neighborhood U_f of f in the product space $\mathbb{R}^{\mathbb{R}}$, if we replace y_x with the corresponding traces $f(\alpha_x)$ of f in the respective factor spaces $Y_{\alpha_1}, \ldots, Y_{\alpha_k}$:

$$U_f = \pi^*_{\alpha_1}[(f(\alpha_1) - \varepsilon_1, f(\alpha_1) + \varepsilon_1)] \cap \cdots$$

$$\cap\, \pi^*_{\alpha_k}[(f(\alpha_k) - \varepsilon_k, f(\alpha_k) + \varepsilon_k)]\,. \quad (5.5a)$$

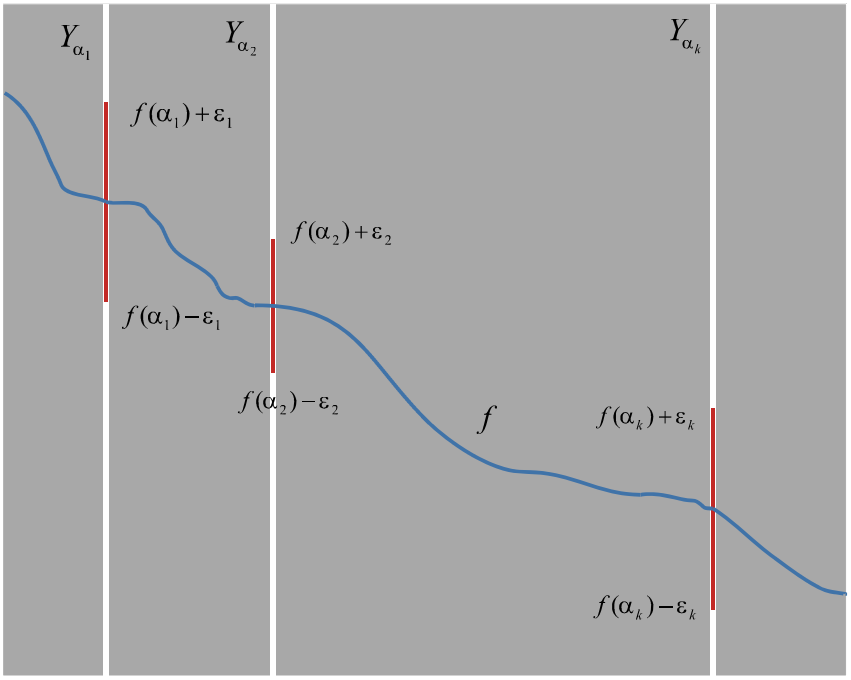


Figure 5.9

So, U_f is the collection of all choice functions from $\mathbb{R}^{\mathbb{R}}$ such that their traces on k fixed factor spaces $Y_{\alpha_1}, \ldots, Y_{\alpha_k}$ must lie within intervals $(f(\alpha_1) - \varepsilon_1, f(\alpha_1) + \varepsilon_1), \ldots, (f(\alpha_1) - \varepsilon_1,\ f(\alpha_1) + \varepsilon_1)$, respectively, and with no further restrictions elsewhere.

PROBLEMS

5.1 For subsets $A \subseteq Y_1$ and $B, C \subseteq Y_2$, prove that

(i) $\mathbf{1}_{A\times B}(x, y) = \mathbf{1}_A(x)\mathbf{1}_B(y), \forall x \in A, y \in B.$

(ii) $(A \times B)^c = (A^c \times Y_2) \cup (Y_1 \times B^c).$

Show the validity of the distributive law with respect to cross product and union, intersection, and difference:

(iii) $A \times (B \cup C) = (A \times B) \cup (A \times C).$

(iv) $A \times (B \cap C) = (A \times B) \cap (A \times C)$.
(v) $A \times (B \backslash C) = (A \times B) \backslash (A \times C)$.

5.2 Prove that if $A_1, A_2 \in Y_1$ and $B_1, B_2 \in Y_2$, then

(i) $(A_1 \cap A_2) \times (B_1 \cap B_2) = (A_1 \times B_1) \cap (A_2 \times B_2)$.

(ii) $(A_1 \times B_1) \backslash (A_2 \times B_2) = [(A_1 \backslash A_2) \times B_1] \cup [A_1 \times (B_1 \backslash B_2)]$.

5.3 Is it true that

$$(A_1 \cap A_2) \times (B_1 \cap B_2) = (A_1 \times B_2) \cap (A_2 \times B_1)?$$

5.4 On the other hand, using the distributive law in Problem 5.1 (iii) it appears that

$$\begin{aligned} &(A_1 \cap A_2) \times (B_1 \cap B_2) \\ &= (A_1 \times B_1) \cap (A_2 \times B_2) \\ &\cap (A_1 \times B_2) \cap (A_2 \times B_1). \end{aligned}$$

Is this correct?

5.5 Find a short expression for

$$(A_1 \cap A_2 \cap A_3) \times (B_1 \cap B_2).$$

5.6 Prove that

(i) $A \times \bigcup_{i \in I} B_i = \bigcup_{i \in I} (A \times B_i)$.

(ii) $\left[\bigcup_{i \in I} A_i\right] \times B = \bigcup_{i \in I} (A_i \times B)$.

(iii) $A \times \bigcap_{i\in I} B_i = \bigcap_{i\in I}(A \times B_i)$.

5.7 Let $Z = \prod_{x\in X} Y_x$, let $A_x \subseteq Y_x$, and let $A = \prod_{x\in X} A_x$, where $A_x = Y_x$ except for finitely many values of the index x, say $1, \ldots, n \in X$. Show that $A = \bigcap_{k=1}^{n} \pi_k^*(A_k)$.

5.8 Let $Z = \prod_{x=1}^{\infty} Y_x$, and let $A_{nx} \subseteq Y_x$, where for each $x = 1, 2, \ldots$, the sequence of sets $\{A_{nx}\}$ is monotone nondecreasing (i.e. $A_{1x} \subseteq A_{2x} \subseteq \cdots$) with

$$\sup\{A_{nx} : n = 1, 2, \ldots\} = \bigcup_{n=1}^{\infty} A_{nx} = Y_x,$$

for $x = 2, 3, \ldots$. Also assume that $A_{11} = A_{21} = A_{31} = \ldots = A_1$. Show that

$$\sup\Big\{ \prod_{x=1}^{\infty} A_{nx} \colon n = 1, 2, \ldots \Big\} = \pi_1^*(A_1).$$

5.9 Let $A = A_1 \times A_2$ and $B = B_1 \times B_2$ be subsets of $Y_1 \times Y_2$ such that $A_i, B_i \subseteq Y_i, i = 1, 2$. Prove the validity of the decomposition

$$A\backslash B = [(A_1\backslash B_1) \times A_2] + [(A_1 \cap B_1) \times (A_2\backslash B_2)].$$

In other words, $A\backslash B$ can be partitioned as two disjoint "rectangles: $(A_1\backslash B_1) \times A_2$ and $(A_1 \cap B_1) \times (A_2\backslash B_2)$.

Show the validity of yet another decomposition of $A\backslash B$:

$$A\backslash B = [A_1 \times (A_2\backslash B_2)] + [(A_1\backslash B_1) \times (A_2 \cap B_2)].$$

5.10 Let $\{A_x \subseteq Y_x;\ x \in X\}$ and $\{Z_y \subseteq Y_y; y \in X\}$ be two family of subsets. Show that

$a)$ $\left(\prod_{x\in X} A_x\right) \cap \left(\prod_{y\in X} B_y\right) = \prod_{x\in X} (A_x \cap B_x).$

$b)$ $\left(\prod_{x\in X} A_x\right) \cup \left(\prod_{y\in X} B_x\right) \supseteq \prod_{x\in X} (A_x \cap B_x).$

5.11 Let $A_1, \ldots, A_m \subseteq Y_1$ and $B_1, \ldots, B_n \subseteq Y_2$. Show that

$$\left[\bigcap_{i=1}^{m} A_i\right] \times \left[\bigcap_{j=1}^{n} B_j\right] = \bigcap_{i=1}^{m} \bigcap_{j=1}^{n} (A_i \times B_j).$$

5.12 Let $\{A_i : i \in I\}$ and $\{B_j : j \in J\}$ be two indexed families of subsets of Y_1 and Y_2, respectively. Show that

$$\left[\bigcap_{i\in I} A_i\right] \times \left[\bigcap_{j\in J} B_j\right] = \bigcap_{i\in I} \bigcap_{j\in J} (A_i \times B_j).$$

5.13 Let $\{S_{i_k}^{(k)} : i_k \in I_k\}$ be an indexed family of subsets of $Y_k, k = 1, \ldots, n$. Show that

$$\prod_{k=1}^{n} \left[\bigcap_{i_k\in I_k} S_{i_k}^{(k)}\right] = \bigcap_{i_1\in I_1} \cdots \bigcap_{i_n\in I_n} \left(\prod_{k=1}^{n} S_{i_k}^{(k)}\right).$$

5.14 Let $m, n \in \mathbb{N}$ and $Y \neq \emptyset$.

$a)$ For $m \leq n$, using a relevant projection map give a bijective function $[\mathcal{P}(Y^m), \mathcal{P}(Y^n), \cdot\,]$.

$b)$ Using a relevant projection map give a bijective function $[\mathcal{P}(Y^n), \mathcal{P}(Y^{\mathbb{R}}), \cdot\,]$.

$c)$ Give a relevant projection map $[Y^n \times Y^{\mathbb{R}}, Y^{\mathbb{R}}, \cdot\,]$.

$d)$ Give a relevant projection map $[Y^{\mathbb{R}} \times Y^{\mathbb{R}}, Y^{\mathbb{R}}, \cdot\,]$.

e) For $A \subseteq X$, using a projection map give a bijective function $[\mathcal{P}(Y^A), \mathcal{P}(Y^X), \cdot\,]$.

5.15 Let $Y_x = \mathbb{R}$, for all $x \in \mathbb{R}$, $A = (0, 2)$.

a) Draw $\pi_A(f)$ for $f(x) = x^2$.

b) Draw $\pi_A^*(\mathcal{A})$ for $\mathcal{A} = (0, 1) \times (0, 1)$.

NEW TERMS:

Descartes, René , 42
Cartesian product of a sequence, 42
Cartesian product of an indexed family of sets, 43
factor space, 43
projection map, 44
$\left[\prod_{x\in X} Y_x\,,\, Y_\alpha\,,\, \pi_\alpha\right]$, αth projection map, 44
αth coordinate of a function, 44
αth factor space, 44
unit cylinder with a base, 46
simple cylinder with a base, 48
rectangle, 49
parallelepiped, 49
rectangular cylinder with a base, 49
$\left[\prod_{x\in X} Y_x\,,\, \prod_{\alpha\in A} Y_\alpha\,,\, \pi_A\right]$, A-projection map, 50
$\pi_A^*(\mathcal{A})$, A-cylinder with base $\mathcal{A}$, 50
neighborhood of a function, 51

6. CARDINALITY

One of the main predicaments in the theory of sets is finding a criterion for their "contents." For finite sets we can count the number of elements they contain and assign them to their respective equivalence classes. We can easily define an equivalence relation in such a class, for example, introducing $\mathbb{C}_n$ as the class of all n-element sets for every $n \in \mathbb{N}_0$. A partial order relation in this class would act as an appropriate comparison among sets from various classes. Sets A and B can be *compared*, in notation $A \preceq B$, if and only if $A \in \mathbb{C}_n$, $B \in \mathbb{C}_s$, and $n \leq s$. Then we could assign to set A the number n and call it the *cardinal number* of A, in notation $\text{card}(A)$ or $|A|$ (which we use interchangeably). Doing this, however, we are up against defining cardinal numbers for infinite sets. A formalism of cardinality belongs to Georg Cantor who was the first to introduce a well-structured concept of "infinity" in his pioneering work done in the 1870s and 1880s.

We present a rather informal version of cardinality sufficient for our goals throughout this book. A curious reader should be referred to special texts on set theory (cf. [Hrba] or [Jech] in the Bibliography Section). We start with comparison of finite sets that will enable us to extrapolate such principles to infinite sets.

6.1 Definitions.

(i) Two sets A and B are said to be *equipotent* if there is a (consistent) bijective function $[A, B, f]$. In this case we denote $\text{card}(A) = |A| = \text{card}(B) = |B|$ (or $A \approx B$) and also say that A and B have *equal cardinality.*

(ii) If there exists an injective function $[A, B, f]$, then we say that the *cardinality of A is less than or equal to the cardinality of B*, in notation $|A| \leq |B|$ or $A \preceq B$. If $|A| \leq |B|$ and $|A| \neq |B|$ we write $|A| < |B|$ or $A \prec B$.

(iii) A *cardinal number* is an equivalence class containing all sets that are " $\approx$ -comparable."

For most cardinal numbers we use the same notation as for regular numbers. For example, rather than using $\mathbb{C}_n$ for the class all n-element sets we just write n.

(iv) Let 0 denote the cardinal number of the empty set Ø (the only representative of this class). Note that 0 is not a number but the class containing Ø. Thus, card(Ø) = 0.

(v) Similarly, the cardinal number n is the equivalence class containing (among other n-element sets) the set $\{1, \dots, n\}$. Therefore, a set A is *finite* if it is equipotent with a set of some cardinal number, say n, such that the integer number n is an element of $\mathbb{N}$, that is, $|A| = |\{1, \dots, n\}| = n$. A set that is not finite is called *infinite.*

(One can easily show that $\mathbb{N}$ is infinite.)

(vi) A set A is said to be *countable* or *denumerable* if it is equipotent with $\mathbb{N}$ and in this case we write $\text{card}(A) = \aleph_0$ (pronounced *aleph naught* or *aleph nought*). A set A is called *at most countable* if $|A| \leq |\mathbb{N}|$ or $A \preceq \mathbb{N}$.

(vii) An infinite set that is not countable is called *uncountable.*

$(viii)$ A set A is said to have the *cardinality of continuum* if it is equipotent with the set $\mathbb{R}$ of real numbers and we write $\text{card}(A) = \mathfrak{c}$. (We show below that $\aleph_0 < \mathfrak{c}$.) □

6.2 Remark. For every set Ω, the property $|A| = |B|$ induces an equivalence relation on the power set $\mathcal{P}(\Omega)$, and $|A| \leq |B|$ induces a partial order on $\mathcal{P}(\Omega)$ (see Problem 6.1). □

6.3 Examples.

(i) If sets A and B have only finitely many elements, then $A \approx B$ if and only if they have the same number of elements.

(ii) In contrast with finite sets, an infinite set can be equipotent with a proper subset of itself. Consider $A = \{1, 3, 5, \dots\} \subseteq \mathbb{N}$ and define $f(n) = 2n - 1, \quad n \in \mathbb{N}$. Then $[\mathbb{N}, A, f]$ is bijective and $\mathbb{N} \approx A$.

(iii) $\mathbb{N}_0 \approx \mathbb{N}_0 \times \mathbb{N}_0$. Indeed, the function

$$f(k, n) = 2^k(2n + 1) - 1$$

is bijective from $\mathbb{N}_0 \times \mathbb{N}_0$ to $\mathbb{N}_0$. Similarly, $\mathbb{N} \approx \mathbb{N} \times \mathbb{N}$.

(iv) Let $\{A_1, A_2, \dots\}$ be a countable family of countable sets. Then its union $A = \bigcup\limits_{n=1}^{\infty} A_n$ is countable. To construct an appropriate bijective map we first represent A as a countable union of disjoint sets. Let

$$B_1 = A_1,\ B_2 = A_2 \backslash A_1, \dots, B_n = A_n \backslash \bigcup_{k=1}^{n-1} A_k \text{ (for } n > 1) \dots .$$

Then, clearly $A = \sum_{n=1}^{\infty} B_n$. Without loss of generality, we assume that each set B_n is countable (some of them can be finite, but it is fair to say that they are at most countable) and, therefore, can be enumerated as $B_n = \{b_{n1}, b_{n2}, \dots\},\ n = 1, 2, \dots$. We can place these sets in the matrix form:

$$\begin{array}{l} B_1 = \{b_{11}, b_{12}, b_{13}, b_{14}, \dots\} \\ B_2 = \{b_{21}, b_{22}, b_{23}, b_{24}, \dots\} \\ B_3 = \{b_{31}, b_{32}, b_{33}, b_{34}, \dots\} \\ B_4 = \{b_{41}, b_{42}, b_{43}, b_{44}, \dots\} \\ \dots\dots\dots\dots\dots\dots\dots\dots \quad . \end{array}$$

Now the desired bijective map $[\mathbb{N}, A, f]$ is formed as follows: $f(1) = b_{11},\ f(2) = b_{12},\ f(3) = b_{21},\ f(4) = b_{31},\ f(5) = b_{22}, f(6) = b_{13}, \ \dots$.

(v) The set $\mathbb{Q}$ of rational numbers is countable, for the function $f(\frac{m}{n}) = (m, n)$ is one-to-one from $\mathbb{Q}$ to $(\mathbb{N} \times \mathbb{N}) \cup \{(0, 0)\}$. The latter is countable by (iii).

(vi) We can show that $\aleph_0 < \mathfrak{c}$. Clearly, $\aleph_0 \leq \mathfrak{c}$. It is sufficient to show that $\mathbb{N} \prec [0, 1]$, for $[0, 1] \approx \mathbb{R}$ (see Problem 6.5). If there is a bijective function $[\mathbb{N}, [0, 1], f]$, then $f(n)$ is of type $0.a_{n_1}a_{n_2}\dots$. Now define the number $0.b_1b_2\dots$ such that $b_i = 3$ if $a_{i_i} \neq 3$ and $b_i = 5$ if $a_{i_i} = 3, i = 1, 2, \dots$. Then the number $b := 0.b_1b_2\dots$ cannot appear among the values of $f(n)$, for it differs from $f(n)$ at the nth place. On the other hand, $b \in [0, 1]$ contradicts the assumption that f is surjective. Thus $\aleph_0 < \mathfrak{c}$. Observe that each rational

number has two representations, for example, 0.1 and 0.0999… . That means we have to be careful about the numbers above that appear "different." □

The following theorem is one of the central results in set theory.

6.4 Theorem (Cantor). $A \prec \mathcal{P}(A)$ *for every set* A.

Proof. The result holds trivially for any finite set A (see Problem 1.4). Specifically, for the empty set, $|\emptyset| = 0$, whereas $|\mathcal{P}(\emptyset)| = 1$. Because $\mathcal{P}(A)$ contains all singletons, it immediately follows that $A \preceq \mathcal{P}(A)$. To show that $|A| \neq |\mathcal{P}(A)|$, we assume that $A \approx \mathcal{P}(A)$ and get a contradiction. By our assumption, there exists a bijective map $f\colon A \to \mathcal{P}(A)$. Then each element a in A is also an element of a subset of A that contains a. In other words, a may belong to $f(a)$ (a subset of A) or may not. We then define $B = \{a \in A\colon a \notin f(a)\}$. B is nonempty, since there exists at least one element $a_0 \in A$ assigned to $\emptyset$. We pick a point $b \in A$ such that $f(b) = B$. By the definition of B, $b \in B \Leftrightarrow b \notin f(b) = B$, and this is a contradiction. □

6.5 Remarks.

(i) In Remark 5.2(iii) we showed that the power set $\mathcal{P}(X)$ of a set X is equipotent with the set $\{0,1\}^X$ of all functions $f\colon X \to \{0,1\}$. Note that 2 is the cardinal number of the set $\{0,1\}$. Thus, we conclude that $|\mathcal{P}(X)| = 2^{|X|}$ (where we set $|B|^{|A|} = |B^A|$). In particular, if $|X| = |\mathbb{N}|$ then $|\mathcal{P}(\mathbb{N})| = 2^{\aleph_0}$. An interesting fact is that $2^{\aleph_0} = \mathfrak{c}$, the proof of which is left for the reader as an exercise (see Problem 6.6).

(ii) The *continuum hypothesis* refines Cantor's Theorem 6.4, asserting that if $\mathfrak{a}$ is an infinite cardinal, then there is no cardinal $\mathfrak{b}$ such that $\mathfrak{a} < \mathfrak{b} < 2^{\mathfrak{a}}$. This was conjectured by Cantor for $\mathfrak{a} = \aleph_0$. In 1900 David Hilbert included the "continuum problem" as Problem #1 in his famous list of open problems in mathematics. In 1940 Kurt Gödel proved that the continuum hypothesis is consistent with (i.e., does not contradict) the axioms of set theory (axiom of existence, axiom of choice, etc.). In 1963 Paul Cohen [Cohe] showed that the continuum hypothesis is independent of the axioms.

(iii) The cardinal number $2^{\mathfrak{c}}$ is called the *hypercontinuum*. For example, the set $\mathcal{P}(\mathbb{R})$ has the hypercontinuum cardinal. □

Supplementary Historical Note. Modern set theory was founded by Georg Cantor, in a sequence of several articles that appeared between 1870 and 1880. One of these articles, "Über eine Eigenschaft des Inbegriffes allen reellen algebraischen Zahlen," appeared in *Crelle's Journal* in 1874, and is said to have given birth to set theory. Georg Cantor was born of Danish parents in St. Petersburg, Russia, in 1845, and lived there until 1856, when his parents moved to Frankfurt, Germany.

Cantor began his university studies at Zürich in 1862. After one semester at Zürich he moved to Berlin University, where he attended lectures of Weierstrass, Kummer, and Kronecker. Leopold Kronecker later became Cantor's main opponent, criticizing his concept of infinity and regarding it as theology and not as mathematics. (As an active theologian, Cantor became a target of attacks by his liberal opponents in Berlin University.) In 1867 Cantor received his Ph.D. (in number theory) from Berlin University. His dream to get a teaching position at Berlin University never came true, primarily due to the opposition of Kronecker.

In 1869 Cantor was appointed to Halle University, where he remained until his retirement in 1913. Cantor died in a mental hospital in Halle in 1918. In 1925 David Hilbert recognized Cantor's concept of infinity. He said, "No one can drive us from the paradise that Cantor created for us." □

PROBLEMS.

6.1 Show the validity of the second statement in Remark 6.2: *For every set* Ω, $|A| \leq |B|$ *induces a partial order on* $\mathcal{P}(\Omega)$.

6.2 Prove the **Schröder – Bernstein theorem**: *If* $A \preceq B$ *and* $B \preceq A$, *then* $A \approx B$.

6.3 (a) Let A be an uncountable set and let $B \subseteq A$ be countable. Show that $A \backslash B$ is uncountable. (b) If $|A| = \mathfrak{c}$ and $|B| \leq \aleph_0$, show that $|A \backslash B| = \mathfrak{c}$.

6.4 Prove that every subset of a countable set is at most countable. (*Hint*: Use the well-ordering principle.)

6.5 Show that $\mathbb{R} \approx [0, 1]$.

6.6 Show that $2^{\aleph_0} = \mathfrak{c}$.

6.7 Let $[X, Y, f]$ be a surjective map. Show that there is a subset of X equipotent with Y.

6.8 Let $[X, Y, f]$ be an surjective map, where Y is countable, and let $f^*(\{y\})$ be a countable set for each $y \in Y$. Must X be countable?

6.9 We call an *algebraic number* any root of a polynomial with integer coefficients. What is the cardinal number of all algebraic numbers?

6.10 Prove the statement. *Every infinite set contains a countable subset.*

6.11 What is the cardinal number of all polynomials whose coefficients are algebraic numbers?

6.12 Show that the set of all finite subsets of $\mathbb{N}$ is countable.

6.13 Show that the set of all finite subsets of a countable set is countable.

6.14 Let $\mathfrak{A} = \{B_x : x \in X\}$ be a countable family (i.e., $X \approx \mathbb{N}$) of countable sets. Let

$$\mathcal{B} = \{B_{x_1} \times \cdots \times B_{x_n} : n \in \mathbb{N}, A = \{x_1, \ldots, x_n\} \subseteq X\}$$

denote the set of all finite Cartesian products from $\mathfrak{A}$. Show that $\mathcal{B}$ is countable.

6.15 Show that if f is a monotone (nondecreasing or nonincreasing) real-valued function on interval $[a, b]$, then the number of discontinuities of f is at most countable.

NEW TERMS:

comparison of sets, 58
cardinal number, 58
Cantor, Georg, 58
equipotent sets, 58
cardinality, 58
finite set, 59
infinite set, 59
countable (denumerable) set, 59
denumerable (countable) set, 59
$\aleph_0$ (aleph naught), 59
at most countable set, 59
uncountable set, 59
continuum, 59
$\mathfrak{c}$, cardinal number of continuum, 59
Cantor's Theorem, 61
continuum hypothesis, 61
Cohen, Paul, 61
hypercontinuum, 62
Cantor, Georg 62
Kronecker, Leopold, 62
Hilbert, David, 62
Schröder – Bernstein theorem, 62
algebraic number, 63

7. BASIC ALGEBRAIC STRUCTURES

Algebra is a mathematical discipline that studies *algebraic structures*. The most rudimentary algebraic operations with natural and positive rational numbers were already encountered in ancient mathematical texts. The famous book, *Arithmetics*, by the Greek Diophantos (of Alexandria) in the third century A.D., has a significant influence on the development of algebraic formalism. The term "algebra" stems from the text *Al-jabr wa'l-mukhabala*, by Muhammad Al-Khwarismi (an Arab mathematician of Persian origin) in the ninth century A.D., which dealt with solution techniques for various problems that reduced to first- and second-order algebraic equations. Not until the end of the fifteenth century, when the common algebraic operations $+$, $-$, $\times$, power, roots, and parentheses were introduced, did one use cumbersome phrases and descriptions of algebraic expressions. François Viéte, by the end of the sixteenth century, was the first to use letters to denote unknowns and parameters.

The algebraic symbolism, as we know it now, has been used only since the middle of the seventeenth century. *Elementary algebra* (which deals with arithmetic operations on real numbers, first- to fourth-order algebraic equations, binomial formula, Diophantine equations) was completed by the middle of the eighteenth century. Leonard Euler's *Introduction to Algebra* was one of the most prominent texts then. In the early nineteenth century, the algebra became furnished with five basic (commutative and associative and distributive) laws with respect to two algebraic operations, $+$ (addition) and $\cdot$ (multiplication). On the strength of Dirichlet's definition of a function, later on, these operations were declared to be binary operations based on the following definition. An *operation* on a set A is a rule that assigns to each ordered subset $A_n \subseteq A$ of n elements a uniquely defined element of the same set A. For $n = 1, 2,$ and $3,$ the operation is called *unary*, *binary*, and *ternary*, respectively.

The *algebraic structures* were formalized in 1830 by the British mathematicians George Peacock in 1830, Duncan Gregory in 1840, and Augustus De Morgan and further refined by the Germans

Hermann Hankel and Hermann Grassman. *Abstract algebra* is regarded as having been born in 1846, when Joseph Liouville published Galois' theory (of solvability of polynomial equations) based on the *group* concept, which began to spread within mathematics. In 1872, the German Felix Klein published a program in which he proposed to formulate all of geometry as the study of invariants under groups of transformations. In 1883, Norwegian mathematician Marius Sophus Lie published his fundamental work on continuous groups of transformations used in studies of continuous functions. Group theory, which is at the heart of contemporary abstract algebra, made prominent contributions to geometry, topology, and even physics in the twentieth century.

In this section, we review some familiar algebraic structures. These provide a basis for analysis, shifting to more abstract settings in the upcoming chapters.

7.1 Definitions.

(i) A set $\mathcal{G}$ with a binary *algebraic operation* $*$ (frequently referred to as *addition* $+$ or *multiplication* $\cdot$) from $\mathcal{G} \times \mathcal{G}$ into $\mathcal{G}$ is called a *semigroup*, in notation $(\mathcal{G}, *)$, if $*$ is associative. (Note that even though $+$ or $\cdot$ may denote addition and multiplication, they need not mean the conventional algebraic operations known for numbers.)

(ii) A semigroup $(\mathcal{G}, *)$ is called a *monoid* if there is an element $\mathbf{1} \in \mathcal{G}$ (called a *two-sided identity*) such that for all $x \in \mathcal{G}$, $x * \mathbf{1} = \mathbf{1} * x = x$.

(iii) A monoid $(\mathcal{G}, *)$ is called a *group*, if for each $x \in \mathcal{G}$, there is a $*$-inverse x' such that $x * x' = x' * x = \mathbf{1}$.

If $*$ is commutative (semigroup, monoid, or group), $(\mathcal{G}, *)$ is called *commutative* or *Abelian*.

If we use for $*$ symbol $+$ or $-$, $(\mathcal{G}, +)$ is referred to as *additive* or $(\mathcal{G}, \cdot)$ *multiplicative*, respectively. It means that in the context of a certain binary operation, $*$ can be identified as $+$ or $\cdot$.

If $(\mathcal{G}, +)$ is additive, the identity element $\mathbf{1}$, denoted by θ, is called *zero*, and the element x' denoted by $-x$ is said to be an *additive inverse* of x.

If $(\mathcal{G}, \cdot)$ is multiplicative, the element $\mathbf{1}$ is called the *unity*. The element x' is denoted by x^{-1} and is said to be a *multiplicative inverse* of x.

(iv) A set $\mathcal{R}$ with addition $+$ and multiplication $\cdot$ from $\mathcal{R} \times \mathcal{R}$ into $\mathcal{R}$ [i.e., a triple $(\mathcal{R}, +, \cdot)$], is called a *ring* if:

$a)$ $(\mathcal{R}, +)$ is an Abelian group,

$b)$ $\cdot$ is associative,

$c)$ $\forall\, a, b, x \in \mathcal{R}$,

$$x \cdot (a + b) = x \cdot a + x \cdot b \quad \text{(called the left distributive law)}$$

$$(a + b) \cdot x = a \cdot x + b \cdot x \quad \text{(called the right distributive law)}.$$

Observe that multiplication $\cdot$ need not be commutative in a ring. However, if this is the case, the *ring* is called *commutative* or *Abelian*. A ring need not have a unity (multiplicative identity) either; consequently, a ring equipped with a unity is called a *ring with unity*. [For instance, the set of all $n \times n$ matrices $\mathcal{M}_{(n,n)}$ is a noncommutative ring with unity (unit matrix).]

(v) Let $(\mathcal{G}, *)$ and $\left(\widetilde{\mathcal{G}}, \widetilde{*}\right)$ be two groups and let $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ (or more precisely, $\left[(\mathcal{G}, *), \left(\widetilde{\mathcal{G}}, \widetilde{*}\right), f\right]$) be a map preserving the algebraic operations $*$ and $\widetilde{*}$, that is, such that

$$f(x * y) = f(x) \,\widetilde{*}\, f(y).$$

Then f is called a (*group*) *homomorphism from* $\mathcal{G}$ *into* $\widetilde{\mathcal{G}}$.

If $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ is bijective then it is called an *isomorphism*. In this case, the groups $(\mathcal{G}, *)$ and $(\widetilde{\mathcal{G}}, \widetilde{*})$ are called *isomorphic*. If $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ is a homomorphism, and in addition, $(\mathcal{G}, *) = (\widetilde{\mathcal{G}}, \widetilde{*})$, then $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ is

said to be an *endomorphism*. If $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ is an endomorphism and an isomorphism then it is called an *automorphism*. □

A homomorphism preserves some (but not all) structural properties of groups, as the following theorem states.

7.2 Theorem. *Let* $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ *be a homomorphism. Then*

(i) *for each* $x \in \mathcal{G}$, $f(x') = [f(x)]'$, *and*

(ii) $f(\mathbf{1}) = \widetilde{\mathbf{1}}$.

(See Problem 7.4.)

7.3 Definition. Let $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ be a homomorphism. Define

$$\mathrm{Ker} f = f^*\big(\{\widetilde{\mathbf{1}}\}\big)$$

and call it the *kernel* of f. □

7.4 Examples.

(i) The space of all continuous functions with operation $+$ forms an Abelian group. The same space is not a group with operation of multiplication.

(ii) All polynomials with operation $+$ form an Abelian group.

(iii) $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, and $((0, \infty), \cdot)$ are Abelian groups, while $(\mathbb{Z}, \cdot)$ is an Abelian monoid.

(iv) The space $\mathbb{C}\backslash\{(0,0)\}$ with the operation complex multiplication $\cdot$ is obviously an Abelian group and $(\mathbb{C}, +, \cdot)$ is a ring.

(v) Let $\mathcal{C}^{(n)}_{[a,b]} := \mathcal{C}^{(n)}([a,b]; \mathbb{R})$ denote the space of all n times continuously differentiable real-valued functions on $[a,b] \subset \mathbb{R}$. Then $(\mathcal{C}^{(n)}_{[a,b]}, +)$ is a commutative group. If $\mathcal{D}^n f$ denotes the nth derivative of a function f, then $\left[\left(\mathcal{C}^{(n)}_{[a,b]}, +\right), \left(\mathcal{C}^{(0)}_{[a,b]}, +\right), \mathcal{D}^n\right]$ is a group homomorphism. Replacing $\mathcal{C}^{(n)}_{[a,b]}$ with the space $\mathscr{P}$ of all polynomials on $[a,b]$, we have $[\mathscr{P}, \mathscr{P}, \mathcal{D}^n]$ as an endomorphism.

(vi) Consider two groups $(\mathbb{R}, +)$ and $((0,\infty), \cdot)$ and the function $f(x) = e^x$. Then, $[\mathbb{R}, \mathbb{R}_+, f]$ is an isomorphism. Indeed, $f(x+y) = f(x) \cdot f(y)$. In addition, $[\mathbb{R}, \mathbb{R}_+, f]$ is bijective.

(vii) Let $\mathcal{F} = \mathcal{F}(X;Y) = Y^X$ be the space of all functions from X to Y. Then, $(\mathcal{F}, \cdot)$ is a multiplicative monoid. For any nonnegative integer n and $f \in \mathcal{F}$, define the unary operation *power* f^n on $\mathcal{F}$ as $f^0 = \mathbf{1}$, $f^{n+1} = f \cdot f^n$. The power has the properties, $f^i \cdot f^k = f^{i+k}$ and $(f^i)^k = f^{ik}$. Note that the power can be defined on an arbitrary multiplicative monoid with the above properties.

$(viii)$ The set $\mathcal{M}^*_{(2,2)}$ of all nonsingular 2×2 matrices is a multiplicative group with the two-sided unity $\mathbf{1} = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, because for each $A \in \mathcal{M}^*_{(2,2)}$, there is a unique two-sided multiplicative inverse A^{-1}.

Now, a function T from $\overline{\mathbb{C}}$ onto $\overline{\mathbb{C}}$ (where $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$) is called a *bilinear transformation* if $T(z) = \frac{az+b}{cz+d}$ if $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$. Let $\mathcal{T}$ denote the set of all bilinear transformations. Then, $(\mathcal{T}, \circ)$, where $\circ$ stands for composition, is a multiplicative group with the two-sided unity $T = \mathbf{1}$ such that $a = d = 1,\ b = c = 0$.

Indeed, firstly we notice that a bilinear transformation T in the form $T(z) = \frac{a_{11}z+a_{12}}{a_{21}z+a_{22}}$ can be associated with a nonsingular matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathcal{M}^*_{(2,2)}$$

through the function f defined as $T = f(A)$. Furthermore, it can be readily shown that the composition $T \circ S$ of T with $S(u) = \frac{b_{11}u+b_{12}}{b_{21}u+b_{22}} = f(B)$, where

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \in \mathcal{M}^*_{(2,2)},$$

can be associated with the matrix product $AB = A \cdot B$. Consequently, T^{-1} stands in association with A^{-1}. In other words, $f(A^{-1}) = T^{-1}$.

In a nutshell, $\left(\mathcal{M}^*_{(2,2)}, \cdot\right)$ and $(\mathcal{T}, \circ)$ are two isomorphic groups, with isomorphism $\left(\mathcal{M}^*_{(2,2)}, \mathcal{T}, f\right)$ such that

$$f(AB) = T \circ S = f(A) \circ f(B).$$

The above discussion agrees with Theorem 7.2 where

(a) for each $A \in \mathcal{M}^*_{(2,2)}$, $f(A^{-1}) = [f(A)]^{-1}$.

(b) $f(I_2) = \mathbf{1}$.

(ix) Let $\mathcal{F} = \{[X, X, f_t];\ t \geq 0\}$ be an indexed family of functions and let $*$ be some binary operation defined on $\mathcal{F}$. $(\mathcal{F}, *)$ is called a *semigroup* (*of functions*) if $f_0 = \mathbf{1}$ and for all $s, t \geq 0$, $f_s * f_t = f_{s+t}$. Obviously, the semigroup $(\mathcal{F}, *)$ is an Abelian monoid.

(x) Let l^1 $(\subseteq \mathbb{R}^{\mathbb{N}})$ be the space of all sequences such that for each $\mathbf{x} = (x_0, x_1, \dots) \in l^1$, $\sum_{n=0}^{\infty} \mid x_n \mid < \infty$. Define the following operation on l^1. For $\mathbf{x}$ and $\mathbf{y}$, let $\mathbf{z} = (z_0, z_1, \dots) = \mathbf{x} * \mathbf{y}$ is such that $z_n = \sum_{k=0}^{n} x_k y_{n-k}$ (called *discrete convolution*). The operation $*$ is commutative and associative and it is closed within l^1 (see Problem 7.12). Obviously, $\mathbf{1} = (1, 0, 0, \dots)$ is the two-sided unity of $(l^1, *)$ and thus $(l^1, *)$ is an Abelian monoid.

(xi) Let $\mathcal{F}_* = \mathcal{F}_*(\mathbb{R}; \mathbb{R})$ denote the space of all bounded real-valued functions. For a function $A \in \mathcal{F}_*$, define

$$u \ \mapsto \ e^{A(u)} = \sum_{n=0}^{\infty} \frac{[A(u)]^k}{k!}$$

in agreement with the unary operation power of Example (vii). Obviously, for each u, the above series converges absolutely, because there is a positive constant M such that

$$\sum_{n=0}^{\infty} \frac{|A(u)|^k}{k!} \ \leq \sum_{n=0}^{\infty} \frac{M^k}{k!} \ = e^M,$$

so that e^A is again an element of $\mathcal{F}_*$. For a fixed A, define the family of functions $f_t = e^{tA}$, $t \geq 0$. From the above definition of e^A it follows that $f_0 = \mathbf{1}$. It is easy to show that $e^{sA}e^{tA} = e^{(s+t)A}$. Indeed,

$$\sum_{i=0}^{n} \frac{s^i A^i}{i!} \sum_{k=0}^{n} \frac{t^k A^k}{k!} = \sum_{r=0}^{n} \sum_{\substack{i+k=r \\ \{0 \leq i,k \leq n\}}} \frac{s^i t^k}{i!k!} A^r$$

$$= \sum_{r=0}^{n} \frac{A^r}{r!} \sum_{i=0}^{r} \frac{s^i t^{r-i} r!}{(r-i)!i!} = \sum_{r=0}^{n} \frac{A^r}{r!}(s+t)^r.$$

The last expression yields $e^{(s+t)A}$ when letting $n \to \infty$. Consequently, $(e^{tA}, \cdot\,)$ is a special case of a semigroup introduced in (ix). This example can be generalized for operators, for instance, square matrices. To discuss such cases rigorously, one would require the concept of the norm of operators. □

7.5 Definitions.

(i) Let $\mathbb{F}$ be a nonempty set with two binary operations, *addition* and *multiplication* (very often, for the elements of $\mathbb{F}$, we drop the conventional multiplication symbol $\cdot$). $(\mathbb{F}, +, \cdot)$ is called a *field* if it is a commutative ring with unity and if for every $\alpha \neq 0$ there is a multiplicative inverse α^{-1}.

In other words, $\mathbb{F}$ is a field if for all $\alpha, \beta, \gamma \in \mathbb{F}$ the following properties hold.

1) (commutative law) $\alpha + \beta = \beta + \alpha$, $\alpha\beta = \beta\alpha$.
2) (associative law) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$, $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.
3) (zero) there is an element $0 \in \mathbb{F}$ such that $\alpha + 0 = \alpha$
4) (additive inverse) there is an element $-\alpha \in \mathbb{F}$ such that $\alpha + (-\alpha) = 0$.
5) (distributive law) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.
6) (unity) there is an element $1 \in \mathbb{F}$ such that $1\alpha = \alpha$.
7) (multiplicative inverse) for every $\alpha \neq 0$, there is $\alpha^{-1} \in \mathbb{F}$ such that $\alpha\alpha^{-1} = 1$.

The elements of a field are called *scalars.*

(ii) Let $\mathbb{F}$ be as above with the exception that $\mathbb{F}$ does not have additive inverses. Then $\mathbb{F}$ is called a *semifield*. We will denote a semifield by $\mathbb{F}_+$.

[The set of all rational numbers, $\mathbb{Q}$, the set of real numbers, $\mathbb{R}$, and the set of all complex numbers, $\mathbb{C}$, are the most common examples of fields. The set of all nonnegative rational or real numbers and the set of complex numbers $z \in \mathbb{C}$ with $\mathrm{Re}(z) \geq 0$, are examples of semifields.]

(iii) A *linear* or *vector space X over a field* $\mathbb{F}$ is a nonempty set with the binary operations addition $(+)$ from $X \times X$ into X and multiplication $(\cdot)$ from $\mathbb{F} \times X$ into X such that

1) $+$ is commutative and associative,
2) there exists an element (called an *origin of* X), $\theta \in X$ such that $0 \cdot x = \theta$, $\forall x \in X$,
3) $1 \cdot x = x$, $\forall\, x \in X$,
4) $\alpha(x + y) = \alpha x + \alpha y$, $(\alpha + \beta)x = \alpha x + \beta x$, $\forall\, \alpha, \beta \in \mathbb{F}$, $\forall\, x, y \in X$,
5) $\alpha(\beta x) = (\alpha\beta)x$, $\forall\, \alpha, \beta \in \mathbb{F}$, $\forall\, x \in X$.

A vector space X over the field $\mathbb{F}$ or semifield $\mathbb{F}_+$ is often denoted by $(X, \mathbb{F}, +, \cdot)$ or $(X, \mathbb{F}_+, +, \cdot)$.

(iv) Elements of X are frequently called *vectors*. If $\mathbb{F} = \mathbb{R}$ then X is called a *real vector space*. If $\mathbb{F} = \mathbb{C}$ then X is called a *complex vector space*. If in (iii) a field $\mathbb{F}$ is replaced by a semifield $\mathbb{F}_+$, then we call X a *semilinear* or *semivector space*.

(v) Any subset of a vector space, which itself is a vector space, is referred to as a *vector* (or *linear*) *subspace* (*subspace* for short) or a *linear manifold*.

(vi) A ring $(\mathcal{A}, +, \cdot)$ is called an *algebra over a field* $\mathbb{F}$ if its additive (Abelian) group $(\mathcal{A}, +)$ is a vector space over $\mathbb{F}$. An algebra over a field $\mathbb{F}$ is denoted by $(\mathcal{A}; \mathbb{F})$. If $(\mathcal{A}; \mathbb{F})$ is an algebra, a pair $(\mathcal{A}'; \mathbb{F}')$ is called a *subalgebra* [of $(\mathcal{A}; \mathbb{F})$], if $\mathcal{A}' \subseteq \mathcal{A}$, $\mathbb{F}' \subseteq \mathbb{F}$, and $(\mathcal{A}'; \mathbb{F}')$ is also an algebra. The above characteristics of commutative rings and rings with unities are hereditary for algebras. □

7.6 Properties of Vector Spaces.

(i) By Definition 7.5(iii), 2) and 3), we have $\theta + x = 0 \cdot x + 1 \cdot x = (0 + 1) \cdot x = x$. Therefore, the origin θ is a zero and, according to Problem 7.1, it is unique.

(ii) For every $x \in X$, there exists $-x$ such that $x + (-x) = \theta$. Indeed, by Definition 7.5(iii), 2) and 4), we have

$$\theta = 0 \cdot x = [1 + (-1)] \cdot x = x + (-1) \cdot x.$$

We call $(-1)\,x$ the *additive inverse of* x and denote it by $-x$. [The reader should not confuse the additive inverse $-\alpha$ of a scalar α previously introduced in Definition 7.5(i) with the additive inverse $-x$ of a vector x first appearing in 7.6(ii).]

(iii) Properties (i) and (ii) imply that a vector space $(X, +)$, with the operation of addition, is an Abelian group.

(iv) $\forall\, \alpha \in \mathbb{F},\ \alpha\theta = \alpha(0 \cdot x) = (\alpha\, 0) \cdot x = 0 \cdot x = \theta.$ □

7.7 Definition (Axioms of a Partially Ordered Vector Space).

(i) A vector space $(X, \mathbb{F}, +, \cdot)$ is said to be *partially ordered*, in notation $(X, \mathbb{F}, +, \cdot, \preceq)$, if it obeys the following axioms.

(A1) $\mathbb{F} = \mathbb{R}$.

(A2) $x \preceq y \Rightarrow x + z \preceq y + z$ for each $z \in X$.

(A3) $\theta \preceq x \Rightarrow \theta \preceq \alpha x$ for each $\alpha \geq 0$.

(ii) From (i) we see that $x \preceq y \Leftrightarrow \theta \preceq y - x$. We may say then that the partial order on X is entirely determined by $X^+ := \{x \in X : \theta \preceq x\}$. The subset X^+ is called the *positive cone of* X. □

7.8 Proposition (Properties of Partially Ordered Vector Spaces). *Let* $(X, \mathbb{R}, +, \cdot, \preceq)$ *be a partially ordered vector space and let* $A, B \subseteq X$. *Then the following properties hold*.

(i) If $x \succeq \theta$, then $-x \preceq \theta$.

(ii) If $x \succeq \theta$ and $\alpha \leq 0$, then $\alpha x \preceq \theta$.

(iii) For $A \subseteq X$ and $x \in X$, $\sup(A + x) = \sup A + x$.

(iv) $\inf(A + x) = \inf A + x$.

(v) $\sup(-A) = -\inf(A)$ and $\inf(-A) = -\sup A$.

(vi) For any $\alpha \in \mathbb{R}_+$, $\alpha \sup A = \sup(\alpha A)$.

(vii) $\sup(A + B) = \sup A + \sup B$.

Proof.

(i) Indeed, by Axiom 2,

$$-x + x \succeq \theta + (-x) \Leftrightarrow \theta \succeq -x.$$

(ii) Indeed,

$\alpha = (-1)(-\alpha)$ and since $-\alpha \geq 0$, by Axiom 3,

$-\alpha x \succeq 0$. But then, by Property (i), $(-1)(-\alpha x) \preceq \theta$.

(iii) Indeed, let $\mathcal{U}(A) = \{y \in X : y \succeq A\}$ and $u_0 = \sup A$. Obviously, for each $a \in A$ and $u \in \mathcal{U}(A)$, $a \preceq u_0 \preceq u$. By Axiom 2, then

$$a + x \preceq u_0 + x \preceq u + x.$$

From the second relation, we have

$$A + x \preceq u_0 + x \preceq \mathcal{U}(A + x).$$

Therefore, $u_0 + x = \sup(A + x)$.

(iv) Indeed, let us denote the set $\mathcal{L}(A) = \{y \in X : y \preceq A\}$ of all lower bounds of A and let $l_\infty = \inf A$. For each $a \in A$ and $l \in \mathcal{L}(A)$, we have that $l \preceq l_\infty \preceq a$. By Axiom 2,

$$l + x \preceq l_\infty + x \preceq a + x$$

or in the form

$$\mathcal{L}(A + x) \preceq l_\infty + x \preceq A + x$$

thereby implying that $l_\infty + x = \inf(A + x)$.

(v) Let $\mathcal{L}(A)$ and $\mathcal{U}(A)$ be the sets of all lower and upper bounds of A and denote $l_\infty = \inf A$ and $u_0 = \sup A$. Then, we have

$$\mathcal{L}(A) \preceq l_\infty \preceq A \preceq u_0 \preceq \mathcal{U}(A).$$

By Property (i),

$$-\mathcal{L}(A) \succeq -l_\infty \succeq -A \succeq -u_0 \succeq -\mathcal{U}(A).$$

Now, $-\mathcal{L}(A) = -\{l \in X : l \preceq A\} = \{-l \in X : -l \succeq -A\}$

$$\Rightarrow -\mathcal{L}(A) = \mathcal{U}(-A).$$

Similarly, $-\mathcal{U}(A) = \mathcal{L}(-A)$. So, we have

$$\mathcal{L}(-A) \preceq -u_0 \preceq -A \preceq -l_\infty \preceq \mathcal{U}(-A).$$

Therefore, $-u_0 = \inf(-A)$ and $-l_\infty = \sup(-A)$.

(vi) Indeed, if $u_0 = \sup A \preceq \mathcal{U}(A)$, we have

$A \preceq u_0 \preceq \mathcal{U}(A)$ or $a \preceq u_0 \preceq u$, for any $a \in A, u \in \mathcal{U}(A)$.

Thus, by Axiom 3, $\alpha A \ni \alpha a \preceq \alpha u_0 \preceq \alpha u \in \mathcal{U}(\alpha A)$ and it yields that $\alpha u_0 = \sup(\alpha A)$.

(vii) Using property (iii) we have

$$\sup A + \sup B = \sup(A + \sup B)$$

$$= \sup\{a + \sup B : a \in A\}$$

$$= \sup\{\sup\{a + b : b \in B\} : a \in A\}$$

$$= \sup\{a + b : a \in A \text{ and } b \in B\} = \sup(A + B). \qquad \square$$

7.9 Definition.

(i) Recall [Definition 4.7 (vi)] that a partially ordered set $(X, \preceq)$ is a lattice if given any two-element subset $\{x, y\} \subseteq X$, the supremum $x \vee y$ and infimum $x \wedge y$ are also elements of X. A partially ordered vector space $(X, \mathbb{R}, +, \cdot, \preceq)$, which is also a lattice, is called a *vector lattice* or *Riesz space*.

(ii) Let X be a semivector space over the semifield $\mathbb{R}_+$, which is a lattice. Then, the space $(X, \mathbb{R}_+, +, \cdot, \preceq)$ is called a *semivector lattice*.

(iii) Let X be a vector lattice. Then $\forall\, x \in X,$ define

$$x^+ := x \vee \theta \text{ (*positive part*)}$$
$$x^- := (-x) \vee \theta \text{ (*negative part*)}$$
$$|x| := x \vee (-x) \text{ (*modulus*)}. \qquad \square$$

7.10 Properties of Vector Lattices.

(i) $(x \vee y) + z = (x + z) \vee (y + z)$. This follows directly from Property 7.8(iii).

(ii) $(x \wedge y) + z = (x + z) \wedge (y + z)$. This follows directly from Property 7.8(iv).

(iii) $x = x^+ - x^-$. Indeed, by (i),

$$(x \vee \theta) - x = \sup((x - x), \theta - x) = (-x) \vee \theta = x^-.$$

(iv) $x^+ + x^- = |x|$. Indeed, by (iii),

$$x^+ + x^- = x + x^- + x^- = x + 2x^-$$

$$= 2[(-x) \vee \theta] + x \text{ [by Property 7.8}(vi)\text{]}$$

$$= (-2x \vee \theta) + x = -x \vee x = |x|. \qquad \square$$

7.11 Examples.

(i) $\{\theta\}$ is a subspace, since by Property 7.6(iii), $\alpha \cdot \theta = \theta$.

(ii) Any field is a vector space over itself.

(iii) $\mathbb{R}^n$ is a real vector space with $\theta = (0, \ldots, 0)$ over $\mathbb{R}$.

(iv) l^1 space, with all real sequences over the field $\mathbb{R}$ whose series are absolutely convergent, is a vector space.

(v) l^p space over the field $\mathbb{C},$ of all sequences such that for each $\boldsymbol{x} = (x_1, x_2, \ldots) \in l^p$, $\Sigma_{n=1}^{\infty} |x_n|^p < \infty,$ where $p \in [1, \infty),$ is a vector space. (See Problem 7.11.)

(vi) $\mathcal{C}_{[a,b]}$ space of all continuous functions on $[a,b]$ is a real vector space.

(vii) $\mathcal{C}^n_{[a,b]}$ space of all n-times differentiable functions on $[a,b]$ is a real vector space.

$(viii)$ $\mathcal{C}^\infty$ space of all infinitely differentiable functions is a complex vector space.

(ix) In Example 7.4 (x), $(l^p \backslash l^0_p \cup \{\theta\}, +, *)$, where $\theta = (0, 0, \ldots)$, is a field, because elements of $l^p \backslash l^0_p$ have multiplicative inverses. $(\mathbb{C}, +, \cdot)$ is another example of a field.

(x) The space $\mathbb{R}^X$ of all real-valued functions on a set X is a commutative algebra over $\mathbb{R}$ with unity. $\mathbb{R}^X$ is also a vector lattice. Indeed, firstly, $(\mathbb{R}^X, +, \cdot)$ is a vector space over the field $\mathbb{R}$. Secondly, a common partial order on $\mathbb{R}^X$ can be defined as $f \preceq g$ if and only if $f(x) \leq g(x)$ for all $x \in X$. Furthermore, $\mathbb{R}^X$ is a lattice. For example, we set $f \vee g = \{\max\{f(x), g(x)\} : x \in X\}$.

(xi) The subspace $\mathcal{F}_*(X;\mathbb{R}) \subseteq \mathbb{R}^X$ of all bounded real-valued functions on a set X is a commutative subalgebra with unity and a vector lattice.

(xii) The subspace $\mathcal{C}(X;\mathbb{R})$ of all continuous functions is also a commutative subalgebra over $\mathbb{R}$ with unity and a vector lattice.

$(xiii)$ The subspace $\mathcal{C}_*(X;\mathbb{R})$ of all bounded continuous functions is a commutative subalgebra of $\mathcal{C}(X;\mathbb{R})$ and a vector lattice.

(xiv) The subspace $\mathcal{C}^n(\mathbb{R};\mathbb{R})$ of all n-times differentiable functions (likewise, the subspace $\mathcal{C}^{(n)}(\mathbb{R};\mathbb{R})$ of n-times continuously differentiable functions) is a commutative subalgebra with unity but not a lattice $(\sup\{x, -x\} = |x| \notin \mathcal{C}^n(\mathbb{R},\mathbb{R}))$.

(xv) The space $\mathcal{C}^\infty(\mathbb{C};\mathbb{C})$ of all infinitely differentiable functions over $\mathbb{C}$ is a commutative algebra with unity but not a lattice.

(xvi) The space $\mathscr{P}$ of all polynomials with real coefficients is a commutative subalgebra over $\mathbb{R}$ with unity but not a lattice.

$(xvii)$ The space $\mathcal{Q}$ of all polynomials with rational coefficients is a commutative subalgebra over the field of rational numbers with unity but not a lattice. □

PROBLEMS

7.1 Show that each monoid has exactly one identity.

7.2 Let $(\mathcal{G}, *)$ be a group. Show that for each two elements $x, y \in \mathcal{G}$, there are $l, r \in \mathcal{G}$, such that $l * x = y$ and $x * r = y$.

7.3 An operation $*$ is called *reducible* if $x * y = x * z$ implies that $y = z$ for all x, y, z. Show that if $(\mathcal{G}, *)$ is a group, then $*$ is reducible. In particular, show that for each $x \in \mathcal{G}$, its inverse is unique.

7.4 Prove Theorem 7.2.

7.5 Let $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ be an isomorphism. Show that $[\widetilde{\mathcal{G}}, \mathcal{G}, f^{-1}]$ is also an isomorphism.

7.6 Show that in Definition 7.3, $\mathbf{1} \in \mathrm{Ker} f$.

7.7 Let $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ be an isomorphism. Find $\mathrm{Ker} f$.

7.8 Let $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ be a mapping such that $\mathcal{G} = \widetilde{\mathcal{G}} = \mathbb{R}$ with operation $+$ and let $f(x) = [x]$ (i.e., the greatest integer less than or equal to x). Is $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ an endomorphism?

7.9 Let $\mathcal{G}$ be the set of all 2×2 real matrices whose determinants equal 1.

$a)$ Show that $(\mathcal{G}, \cdot)$ (with the usual matrix multiplication) is a group.

b) Let B be any 2×2 nonsingular matrix. Define the map $[\mathcal{G}, \mathcal{G}, f]$ such that $f(A) = B^{-1}AB$. Show that $[\mathcal{G}, \mathcal{G}, f]$ is an automorphism.

7.10 Show that $\forall a, b \geq 0$ and $p \in [1, \infty)$, $(a+b)^p \leq 2^{p-1}(a^p + b^p)$. [*Hint*: For $p > 1$, work with the auxiliary function $f(x) = (a+x)^p - 2^{p-1}(a^p + x^p)$, $x \geq 0$.]

7.11 Show that l^p (defined in Example 7.11 (v)) is a vector space over $\mathbb{C}$; specifically show that $x, y \in l^p \Rightarrow x + y \in l^p$.

7.12 Show that the operation $*$ in Example 7.4 (x) is commutative and associative and it is closed in l^1, that is, show that if $\mathbf{x}, \mathbf{y} \in l^1$, then $\mathbf{x} * \mathbf{y} \in l^1$.

7.13 Is $(\mathcal{T}, +, \circ)$ [where $(\mathcal{T}, \circ)$ is defined in Example 7.4 $(viii)$] a ring?

7.14 Let S be a subset of $\mathbb{C}$. Argue for what cases S is a subspace of $\mathbb{C}$ over $\mathbb{R}$:

a) S is a closed unit disc centered at zero; that is, $S = \{z \in \mathbb{C} : |z| \leq 1\}$.

b) $S = \{z \in \mathbb{C}: \{|\text{Re}(z)| \leq 1\} \times \{|\text{Im}(z)| \leq 1\}\}$.

c) $S = \{z \in \mathbb{C}: \{\text{Im}(z) = 0\} \times \{|\text{Im}(z)| \leq 1\}\}$.

d) $S = \{z \in \mathbb{C} : \text{Im}(z) \geq 0 \text{ and } \text{Re}(z) \geq 0\} \cup \{z \in \mathbb{C} : \text{Im}(z) \leq 0 \text{ and } \text{Re}(z) \leq 0\}$.

7.15 In the condition of Example 7.4(x), let $\mathbf{x} = (x_0, x_1, \dots) \in l^1$ such that $x_0 \neq 0$. Show that for each $\mathbf{x}$ with $x_0 \neq 0$, there is a unique element $\mathbf{x}^{-1}$ with respect to $*$ such that $\mathbf{x} * \mathbf{x}^{-1} = \mathbf{1}$. Is $\mathbf{x}^{-1}$ the two-sided inverse of $\mathbf{x}$? If this is the case, show that $(l^1 \backslash l_0^1, *)$ is an Abelian group, where l_0^1 denotes the subset of all elements $\mathbf{x} \in l^1$ with $x_0 = 0$. Is the triple $(l^1, +, *)$ an Abelian ring with unity?

7.16 Let $(X, \mathbb{R}, +, \cdot, \preceq)$ be a Riesz space and let $A, B \subseteq X$ such that $\inf A, \inf B, \inf(A+B) \in X$. Prove the following equations.

(i) $\inf(A+B) = \inf A + \inf B$.

(ii) For any $\alpha \geq 0$, $\alpha \inf A = \inf(\alpha A)$.

(iii) For any $\alpha \leq 0$, the following hold: $\alpha \sup A = \inf(\alpha A)$ and $\alpha \inf A = \sup(\alpha A)$.

7.17 Show the validity of the following equations in a Riesz space $(X, \mathbb{R}, +, \cdot, \preceq)$.

(i) $2(x \vee y) = x + y + |x - y|$.

(ii) $2(x \wedge y) = x + y - |x - y|$.

(iii) $x + y = (x \vee y) + (x \wedge y)$.

(iv) $x^+ \vee x^- = |x|$.

(v) $x^+ \wedge x^- = \theta$.

(vi) $|\alpha x| = |\alpha||x|$.

(vii) $|x + y| + |x - y| = 2(|x| \vee |y|)$.

7.18 Let $(X, \mathbb{R}, +, \cdot, \preceq)$ be a Riesz space. Prove the following inequalities.

(i) We say that $A \preceq u$ if $u \in \mathcal{U}(A)$. Let $A \preceq u$ and $B \preceq u$, then $\sup(A \cup B) \preceq u$.

(ii) If $x \preceq z$ and $y \preceq z$, then $x \vee y \preceq z$.

(iii) If $x \preceq z$ and $-x \preceq z$, then $|x| \preceq z$.

(iv) The inequality $-z \preceq x \preceq z$ is equivalent to $|x| \preceq z$.

(v) The triangle inequality: $|x+y| \preceq |x| + |y|$.

7.19 Let $(X, \mathbb{R}, +, \cdot, \preceq)$ be a Riesz space. If $x \in X$ and $\alpha \in \mathbb{R}$, prove the following:

(i) $(\alpha x)^+ = \alpha x^+$ and $(\alpha x)^- = \alpha x^-$ if $\alpha \geq 0$.

(ii) $(\alpha x)^+ = -\alpha x^-$ and $(ax)^- = -\alpha x^+$ if $\alpha \leq 0$.

(iii) $(\alpha x)^+ - (\alpha x)^- = \alpha x$, for all $a \in \mathbb{R}$.

7.20 Show that the positive cone X^+ has the following properties:

$X^+ \cap (-X^+) = \{\theta\}$.

$X^+ + X^+ \subseteq X^+$.

$\alpha X^+ \subseteq X^+$, for all $\alpha \geq 0$.

NEW TERMS:

algebra, 66
algebraic structures, 66
Al-Khwarismi, Muhammad, 66
Peacock, George, 66
Gregory, Duncan 6,6,
De Morgan, Augustus, 66
Hankel, Hermann, 67
Grassman, Hermann, 67
abstract algebra, 67
Liouville, Joseph, 67
Klein, Felix, 67
Lie, Marius, 67
algebraic operation, 67
semigroup, 67
associative algebraic operation, 67
monoid, 67
two-sided identity, 67
group, 67
commutative algebraic operation, 67
Abelian algebraic operation, 67
Abelian group, 67
Abelian semigroup, 67
Abelian monoid, 67
additive algebraic operation, 67
multiplicative algebraic operation, 67
identity element, 68
zero, 68
inverse, 68
additive inverse, 68
unity, 68
multiplicative inverse, 68
ring, 68
left distributive law, 68
right distributive law, 68
commutative ring, 68

Abelian ring, 68
ring with unity, 68
group homomorphism, 68
group isomorphism, 68
group endomorphism, 69
group automorphism, 69
kernel, 69
space $\mathcal{C}^{(n)}_{[a,b]}$ of all n times differentiable functions, 69
power, 70
bilinear transformation, 70
semigroup of functions, 71
l^1 space, 71
discrete convolution, 71
field, 72
scalar, 72
semifield, 73
linear (vector) space over a field, 73
vector (linear) space over a field, 73
origin, 73
vector, 73
real vector space, 73
complex vector space, 73
semivector space, 73
subspace, 73
linear manifold, 73
algebra over a field, 73
subalgebra, 73
additive inverse in vector space, 74
$(X, \mathbb{R}, +, \cdot, \preceq)$, partially ordered vector space, 74
positive cone, 74
vector lattice, 76
Riesz space, 76
semivector lattice, 77
x^+, positive part, 77
x^-, negative part, 77
$|x|$, modulus, 77

8. MORE ABOUT VECTOR SPACES

In this section we study vector spaces more systematically in response to the needs of forthcoming topics on normed vector spaces (Chapter 2) and Hilbert spaces [Dsh2]. Most of these materials can be found in linear algebra textbooks.

8.1 Definitions.

(i) Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$. Let $\{v_1, \dots, v_n\}$ be a finite subset of $\mathcal{V}$ and let $\{\alpha_1, \dots, \alpha_n\} \subseteq \mathbb{F}$. The vector

$$\alpha_1 v_1 + \cdots + \alpha_n v_n \in \mathcal{V}$$

is called a *linear combination* of $\{v_1, \dots, v_n\}$. If a linear combination $\alpha_1 v_1 + \cdots + \alpha_n v_n = \theta$ and it implies that $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, then the set $\{v_1, \dots, v_n\} \subseteq \mathcal{V}$ is said to be *linearly independent*. We set by definition the empty set Ø as linearly independent. If a finite subset $\{v_1, \dots, v_n\} \subseteq \mathcal{V}$ is not linearly independent then it is called linearly dependent. An arbitrary subset A of $\mathcal{V}$ is called *linearly independent* if each of its finite subsets is linearly independent. If A has at least one finite linearly dependent subset, then A is said to be linearly dependent.

(ii) The intersection of all subspaces of a vector space $\mathcal{V}$ that contain a certain subset $A \subseteq \mathcal{V}$ is called the *linear span* (or *linear hull*) *of* A and it is denoted by span$A$.

Observe that a linear span is a subspace of $\mathcal{V}$. Indeed, if $x, y \in$ span$A$, then x, y are elements of all subspaces containing A. Obviously, $\alpha x + \beta y$ is an element of all subspaces containing A and thus it belongs to span$A$. All other properties of span$A$ as a vector space can be shown similarly. Obviously, a linear span is the minimal subspace containing set A. We refer to span$A$ as the *subspace generated* (or *spanned*) *by* A. Any subspace $\mathcal{S}$ is a trivial example of its linear span, that is, span$\mathcal{S} = \mathcal{S}$.

(iii) A subset $H \subseteq \mathcal{V}$ is called a *Hamel basis* (or just *basis*) *for* $\mathcal{V}$ if

(a) H is linearly independent,

(b) span$H = \mathcal{V}$. □

8.2 Remarks. Let $\mathcal{V}$ be a vector space.

(i) It is clear that span$\emptyset = \{\theta\}$, because $\{\theta\}$ is the minimal subspace containing the empty set $\emptyset$.

(ii) If $A \neq \emptyset$, then span$A$ equals the set A_0 of all linear combinations of elements of A. Indeed, A_0 is a subspace containing A and therefore, span$A \subseteq A_0$. Let $v_0 \in A_0$. Then $v_0 = \alpha_1 x_1 + \cdots + \alpha_n x_n$ such that $x_i \in A \subseteq$ span$A$. On the other hand, because span$A$ is a subspace itself and because $x_i \in$ span$A$, it must contain v_0. Therefore, $A_0 \subseteq$ span$A$ implying that $A_0 =$ span$A$.

(iii) If H is a Hamel basis for $\mathcal{V}$, then Remark (ii) implies that each vector of $\mathcal{V}$ can be represented as a linear combination of elements of H. We can also convert this statement; namely, if H is a subset of $\mathcal{V}$ such that any vector of $\mathcal{V}$ is a linear combination of elements of H, then H is a Hamel basis. It is easy to see that any such representation is unique relative to one particular Hamel basis. Indeed, let $v \in \mathcal{V}$ and suppose we have two representations of v such that $v = \sum_{i=1}^{n} \alpha_i h_i = \sum_{k=1}^{m} \beta_k h'_k$. Then,

$$\sum_{i=1}^{n} \alpha_i h_i + \sum_{k=1}^{m} (-\beta_k) h'_k = \theta.$$

If h_i's and h'_k's are all distinct, then due to linear independence of H, $\alpha_i = \beta_k = 0$. Even if for some i and k, $h_i = h'_k$'s, again by linear independence of H it follows that $\alpha_i = \beta_k$. □

8.3 Example. Let $\mathcal{V} = \mathbb{C}^n := \mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}$ (n times) be the n-dimensional complex space. It is a vector space over the field $\mathbb{C}$. Denote by $H := \{e_1 \ldots, e_n\}$ the set of unit vectors $e_i := (0, \ldots, 0, 1, 0, \ldots, 0)$ where 1 is the ith entry. This set is linearly independent. Let $\zeta \in \mathbb{C}^n$ be an arbitrary vector with coordinates $(\zeta_1, \ldots, \zeta_n)$ where $\zeta_j = x_j + iy_j$. Then, by choosing $\alpha_j = \zeta_j$ we can represent ζ as $\zeta = \sum_{j=1}^{n} \alpha_j e_j$. Therefore, H is a Hamel basis for $\mathbb{C}^n$. Observe that H is not a unique Hamel basis for $\mathcal{V}$. For example, in the above definition of e_i we can replace the ith element 1 with -1 and do so with some other e_j's . □

8.4 Theorem. *Any two Hamel bases in a vector space $\mathcal{V}$ have the same cardinality*. □

For finite Hamel bases, the proof to Theorem 8.4 is available in most linear algebra text books. For the general case, a proof can be found in [Hew], Theorem 4.58, p. 30.

8.5 Remark. As we have seen, a vector space can have more than one Hamel base. However, it can be proved that all of them have the same cardinal number [card(H)]. In other words, if H_1 and H_2 are two Hamel bases for $\mathcal{V}$, then there is a bijective map $f : H_1 \to H_2$, that is, H_1 and H_2 are equivalent sets. Specifically, if H_1 is finite, then H_2 is also finite and they have the same number of elements, that is, card(H_1) = card(H_2). [Recall that $\aleph_0 :=$ card$(\mathbb{N})$ and $\mathfrak{c} :=$ card$(\mathbb{R})$ are the most frequently used infinite cardinals.] Thus, each vector space $\mathcal{V}$ is characterized by its "dimension," in notation dim$\mathcal{V}$, that is by the cardinal number of one of its Hamel bases. Correspondingly, we distinguish *finite-* and *infinite-dimensional* vector spaces. For instance, in the above example, dim$\mathbb{C}^n = n$. □

8.6 Definitions.

(i) Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over a field $\mathbb{F}$. An $f : \mathcal{V} \to \mathcal{W}$ is called a *linear map* (or *linear operator*) if for all $x, y \in \mathcal{V}$ and $\alpha \in \mathbb{F}$,

(a) $f(x + y) = f(x) \mathbin{\widetilde{+}} f(y)$ (where $\mathbin{\widetilde{+}}$ is the operation of addition in $\mathcal{W}$)

(b) $f(\alpha \cdot x) = \alpha \cdot f(x)$.

(ii) A linear map $f : \mathcal{V} \to \mathbb{F}$ (where $\mathcal{V}$ is a vector space over field $\mathbb{F}$) is called a *linear functional*.

(iii) Replacing field $\mathbb{F}$ in (i) and (ii) by a semifield $\mathbb{F}_+$, we have the notions of a *semilinear map* and a *semilinear functional*, respectively.

(iv) Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over a field $\mathbb{F}$. If the linear map in (i) is bijective, then we call it an *isomorphism*. [For a more general notion of an isomorphism see Definition 7.1(v).] The vector spaces $\mathcal{V}$ and $\mathcal{W}$ are called *isomorphic*. □

8.7 Remarks.

(i) Property (a) in Definition 8.6 (i) makes a linear map $[\mathcal{V}, \mathcal{W}, f]$ a homomorphism from $(\mathcal{V}, +)$ into $(\mathcal{W}, \widetilde{+})$. By Theorem 7.2,

$$\forall\, x \in \mathcal{V},\ f(-x) = -f(x) \text{ and}$$
$$f(\theta_{\mathcal{V}}) = \theta_{\mathcal{W}},$$

which we arrived at without the utility of property (b).

(ii) According to Definition 7.3, $f^*(\theta_{\mathcal{W}}) = \mathrm{Ker} f$ (often referred to as the *null space of* f). By Problem 7.6, $\theta_{\mathcal{V}} \in \mathrm{Ker} f$. □

8.8 Proposition. *If* $[\mathcal{V}, \mathcal{W}, f]$ *is a linear map, then*

(i) *the subset* $\mathrm{Ker} f \subseteq \mathcal{V}$ *is a linear manifold.*

(ii) *the subset* $f_*(\mathcal{V}) \subseteq \mathcal{W}$ *is a linear manifold.* □

(See Problems 8.1 and 8.2.)

8.9 Examples.

(i) Let l_c be the set of all real-valued sequences with all but finitely many zero elements (called the sequences with *finite support*). l_c is a vector space over $\mathbb{R}$ with the usual operations of addition and multiplication by scalars. Clearly, l_c is isomorphic to the space $\mathscr{P}$ of all polynomials with real coefficients.

(ii) If $\mathcal{V}$ is a vector space over a field $\mathbb{F}$, with a finite Hamel basis $H = \{h_1, \ldots, h_n\}$, then each vector $v \in \mathcal{V}$ can be uniquely represented as $v = \alpha_1 h_1 + \alpha_2 h_2 + \cdots + \alpha_n h_n$. Hence, each vector $v \in \mathcal{V}$ can be associated with a unique vector $f(v) = (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ and conversely, each vector $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ can be associated with a unique vector $f^{-1}(\alpha_1, \ldots, \alpha_n) = v \in \mathcal{V}$. Observe that $\mathbb{F}^n$ is a vector space over $\mathbb{F}$. In addition, $\alpha x + \beta y = (\alpha\alpha_1 + \beta\beta_1)h_1 + \cdots + (\alpha\alpha_n + \beta\beta_n)h_n$, which implies that $\mathcal{V}$ and $\mathbb{F}^n$ are isomorphic. This is why we often prefer to use the term "vector space" over the term "linear space." □

Specifically, if $\mathbb{F} = \mathbb{C}$, the above considerations can be summarized as the following proposition.

8.10 Proposition. *Any n-dimensional vector space over the field $\mathbb{C}$ is isomorphic to $\mathbb{C}^n$.* □

The following theorem establishes the existence of a Hamel basis in a vector space.

8.11 Theorem. *Each vector space $\mathcal{V}$ with two or more elements over a field $\mathbb{F}$ has a Hamel basis.*

Proof. Let $\mathcal{F}$ be the family of all linearly independent subsets in $\mathcal{V}$. Clearly, $\mathcal{F} \neq \emptyset$ due to our assumption that card $\mathcal{V} \geq 2$. Obviously, $\mathcal{F}$ satisfies Tukey's Lemma 4.14 according to which $\mathcal{F}$ has at least one maximal element in the sense of partial order $\subseteq$. □

Now, with the aid of Theorem 8.11 we can get a very useful criterion for Hamel bases.

8.12 Theorem. *H is a Hamel basis for a vector space $\mathcal{V}$ over a field $\mathbb{F}$ if and only if H is a maximal linearly independent set.*

Proof. (i) Let H be a Hamel basis and let $x \in \mathcal{V}$ such that $x \neq \theta$. Then, x can be represented as a finite linear combination

$$x = \sum_{i=1}^{n} \alpha_i h_i \ (\alpha_i \in \mathbb{F}, h_i \in H).$$

Thus, $\{x, h_1, \ldots, h_n\}$ is a linearly dependent subset, while $H \cup \{x\}$ is linearly dependent, evidently for any $x \in \mathcal{V} \backslash H$. Therefore, H is indeed a maximal linearly independent set in $\mathcal{V}$.

(ii) Now suppose H is a maximal linearly independent set in $\mathcal{V}$. We will prove that H is a Hamel basis. Let $\theta \neq h_0 \in \mathcal{V} \backslash H$. Then, $H \cup \{h_0\}$ is linearly dependent. Therefore, there is a linear combination $\sum_{i=0}^{n} \alpha_i h_i = \theta$ such that $h_1, \ldots, h_n \in H$ and that not all α_i's are zeros. If $\alpha_0 = 0,$ then $h_1, \ldots, h_n$ are linearly dependent, which is impossible. Thus,

$$h_0 = -\sum_{i=1}^{n} \frac{\alpha_i}{\alpha_0} h_i.$$

Here not all $\alpha_i = 0,$ because $h_0 \neq \theta$ by our assumption. If $h_0 \in H$, then $h_0 = 1 \cdot h_0$. Consequently, $h_0 \in \text{span}H$ and thus $\mathcal{V} \subseteq \text{span}H$. So, H is a Hamel basis. □

And one more useful result.

8.13 Proposition. *Let A be a linearly independent subset of a vector space X. Then, there is a Hamel basis H for X such that $A \subseteq H$.*

Proof. If there is no Hamel basis H for X such that $A \subseteq H$, then set A is a maximal linearly independent set in X and thus is a Hamel basis itself. □

PROBLEMS

8.1 Prove the statement in Proposition $8.8(i)$ that *the subset* $\text{Ker} f \subseteq \mathcal{V}$ *is a linear manifold*.

8.2 Prove the statement in Proposition $8.8(ii)$ that *the subset* $f_*(\mathcal{V}) \subseteq \mathcal{W}$ *is a linear manifold.*

8.3 Let X be a finite-dimensional vector space and let $Y \subseteq X$ be a linear manifold. Show that $\dim Y \leq \dim X$.

8.4 Let X be a finite-dimensional vector space and let $Y \subseteq X$ be a linear manifold such that $\dim Y = \dim X$. Show that $Y = X$.

8.5 Let X be a vector space and $A, B \subseteq X$. Define $A + B = \{x + y : x \in A, y \in B\}$. Prove that if A and B are linear manifolds, then $A + B$ is a linear manifold.

8.6 Let X be a vector space and let $Y, Z \subseteq X$ be two linear manifolds. Prove the statement. *For each $x \in Y + Z$, there is a unique decomposition of x by vectors $y \in Y$ and $z \in Z$ such that $x = y + z$ if and only if $Y \cap Z = \{\theta\}$.* (Two such manifolds are said to be *linearly disjoint.*)

8.7 Let X be a finite-dimensional vector space and $Y, Z \subseteq X$ be two linearly disjoint manifolds. (See Problem 8.6.) Then,

$Y + Z$ is called a *direct sum* of Y and Z and it is denoted by $Y \oplus Z$. Prove that if $X = Y \oplus Z$, then $\dim X = \dim Y + \dim Z$.

8.8 Show that the set $\mathcal{C}^{\infty}_{[a,b]}$ of all infinitely differentiable functions on $[a, n]$ is not finite-dimensional.

8.9 Show that the differential operator $\frac{d^n}{dx^n} : \mathcal{C}^n_{[a,b]} \rightarrow \mathcal{C}_{[a,b]}$ is linear with respect to $\mathbb{R}$.

8.10 Let A be an $n \times m$ matrix. Show that A: $\mathbb{R}^m \rightarrow \mathbb{R}^n$ is a linear operator with respect to $\mathbb{R}$.

NEW TERMS:

linear combination, 85
linear independence, 85
linearly dependent finite set, 85
linear span (hull), 85
linear hull (span), 85
subspace spanned by a set, 85
Hamel basis for a vector space, 85
dimension of vector space, 87
finite dimensional vector space, 87
infinite dimensional vector space, 87
linear map, 87
linear operator, 87
linear functional, 87
semilinear map, 87
semilinear functional, 87
isomorphism, 87
isomorphic vector spaces, 87
null space, 88
l_c space, 88
sequence with finite support, 88
maximal linearly independent set, 89
linearly disjoint manifolds, 90
direct sum, 91

Chapter 2

Analysis of Metric Spaces

Metric spaces were introduced and studied by the French mathematician, Maurice René Fréchet (in his doctoral dissertation published in 1906), and developed later by the German Felix Hausdorff (in his 1914 book *Grundzüge der Mengenlehre*). It was apparent that to the end of the nineteenth century the mathematical world (partly inspired by Cantor's fundamental work in set theory) was eager to structure more general sets than conventional $\mathbb{R}^n$. On the other hand, the needs of complex analysis and the rash development of differential equations accelerated this process. Typical examples are uniform convergence in function spaces, approximation of continuous functions by polynomials and the Riemann mapping theorem. After 1920, the theory of metric spaces, especially fundamental work on normed spaces and their applications to functional analysis, was further developed by the Pole Stefan Banach and his school. Paying tribute to their achievements and of their fellow countrymen followers, an important subclass of metric spaces was named "Polish." A series of studies of metric spaces was further undertaken in the late 1920s by the Russian school of analysis. At this time, metric spaces have become generalized to topological spaces.

In this chapter we introduce the main principles of metric spaces and their special case: normed vector spaces. This part of analysis traditionally precedes the more general theory of topology and functional analysis.

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0_2, © Springer Science+Business Media, LLC 2013

1. DEFINITIONS AND NOTATIONS

The concept of "metric" (measuring distances in space) is at the root of mathematical (geometric) thinking. Starting with that concept we show how the concepts of limits of sequences and continuity of functions can be extended by metrization of spaces more general than Euclidean spaces introduced in calculus. Recall that a point x is a *limit of a sequence* $\{x_n\}$ if all terms of the sequence numbered with $k, k+1, \ldots$ for some k are sufficiently "close" to x. The closeness of these points to x was defined in terms of the Euclidean distance $|x - x_k|$, which determined the specific structure employed on $\mathbb{R}$. In many applications, an underlying carrier is often more general than $\mathbb{R}$ or even $\mathbb{R}^n$. So, the question arises, "how do we build the analysis in the general space?" Since the distance was crucial in the formation of analysis on the real line, we therefore introduce this notion also for the general space, emphasizing the main properties of the distance we have had experience with. Once a distance (or *metric*) between any two points of a given set is defined, the set becomes "well-structured" or *metrized*, and then is ranked as a *space*, more precisely, a *metric space*.

1.1 Definitions.

(i) Let X be a nonempty set. A *metric* d (or *distance*) on X is any nonnegative function $d : X^2 \to \mathbb{R}_+$ such that:

(a) $\forall x, y \in X,\ d(x, y) = 0 \Leftrightarrow x = y.$
(b) $\forall x, y \in X,\ d(x, y) = d(y, x).$
(c) $\forall x, y, z \in X, d(x, y) \leq d(x, z) + d(z, y)$ (*triangle inequality*).

The pair (X, d) is called a *metric space.* As usual, we refer to set X as a *carrier*. Sometimes, for brevity, carrier X itself is called the *metric space*.

(ii) If for $x, y \in X,\ x = y$ implies $d(x, y) = 0$, but the converse does not hold [i.e., $d(x, y) = 0$ does not yield $x = y$], and (b) and (c) hold, then d is called a *pseudo-metric.* Correspondingly, the pair (X, d) is called a *pseudo-metric space.*

Any pseudo-metric can be "made" a metric by introducing the equivalence classes generated by d in such a way that x and y will belong to one and the same class whenever $d(x, y) = 0$. The corresponding metric space will then turn to $(X|\mathrm{E}_d, \boldsymbol{d})$, where $X|\mathrm{E}_d$ is the quotient set of X modulo E_d and $\boldsymbol{d}$ is the reducer or d (as per Theorem 4.4, Chapter 1). □

1.2 Remark. By the triangle inequality we have

$$d(x, y) - d(z,y) \leq d(x, z), \tag{1.2a}$$

which holds for all $x, y, z \in X$. Then, interchanging x and z in the last inequality we arrive at

$$d(x, y) - d(z, y) \geq -d(x, z). \tag{1.2b}$$

Inequalities (1.2a) and (1.2b) yield

$$|d(x, y) - d(y, z)| \leq d(x, z), \ \forall x, y, z \in X. \tag{1.2c}$$

□

Let $Y \subseteq X$. Then the pair (Y, d) is also a metric space, called a *subspace of* (X, d).

1.3 Examples (of metric spaces).

(i) The *discrete metric* is defined on a nonempty set X as

$$d(x, y) = \begin{cases} 1, & x \neq y \\ 0, & x = y\,. \end{cases}$$

The triangle inequality does not hold if and only if $d(x, y) = 1$ and $d(x, z) = d(z, y) = 0$. However, the latter can only be possible when $x = z = y$. Hence, $d(x, y)$ cannot equal 1.

(ii) Let $X = (0, \infty)$ and $d(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right|$. The triangle inequality follows from

$$d(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right| = \left|\frac{1}{x} - \frac{1}{z} + \frac{1}{z} - \frac{1}{y}\right|$$

$$\leq \left|\tfrac{1}{x} - \tfrac{1}{z}\right| + \left|\tfrac{1}{z} - \tfrac{1}{y}\right| = d(x,z) + d(z,y)\,.$$

(iii) Let X consist of all sequences $\{x_n\} \subset \mathbb{R}$. Such a carrier X is denoted by $\mathbb{R}^{\mathbb{N}}$. Recall that a subset of $\mathbb{R}^{\mathbb{N}}$ is the l^1 *space* if it contains only absolutely convergent sequences, i.e., those with

$$\textstyle\sum_{n=1}^{\infty} \mid x_n \mid \; < \infty.$$

Let us define the function d on l^1 as $d(x,y) = \sum_{n=1}^{\infty} \mid x_n - y_n \mid$. Then

$$\begin{aligned} d(x,y) &= \sum_{n=1}^{\infty} \mid x_n - z_n + z_n - y_n \mid \\ &\leq \sum_{n=1}^{\infty} \mid x_n - z_n \mid \; + \sum_{n=1}^{\infty} \mid z_n - y_n \mid \\ &= d(x,z) + d(z,y). \end{aligned}$$

Thus, d is a metric on l^1, because the other properties of d as a metric are obvious.

(iv) Let $\mathcal{C}_{[a,b]}$ denote the set of all continuous functions on interval $[a,b] \subseteq \mathbb{R}$. Let us define

$$d_\infty(x,y) = \sup\{|x(t) - y(t)| : t \in [a,b]\},$$

called the *supremum metric*. Because any continuous function on a closed and bounded interval assumes maximum and minimum values, the definition of d makes sense. The inequalities

$$\begin{aligned} |x(t) - y(t)| &\leq |x(t) - z(t)| + |z(t) - y(t)| \\ &\leq \sup|x(t) - z(t)| + \sup|z(t) - y(t)| \end{aligned}$$

hold for all $t \in [a,b]$. Therefore we have

$$\sup|x(t) - y(t)| \leq \sup|x(t) - z(t)| + \sup|z(t) - y(t)|,$$

which is exactly the triangle inequality. Property (a) of a metric obviously holds. Hence d is a metric on $\mathcal{C}_{[a,b]}$.

(v) Now, define another metric on $\mathcal{C}_{[a,b]}$:

$$d(x,y) = \int_a^b |x(t) - y(t)|dt.$$

It is easy to see that $d(x,y) = 0$ if and only if $x(t) = y(t)$ for all $t \in [a,b]$. (See Problem 1.11). The triangle inequality is obvious. □

PROBLEMS

1.1 Let $X = \mathbb{R}$ and $d(x,y) = \sin^2(x-y)$. Is (X,d) a metric space?

1.2 Let $X = \mathbb{R}$ and $d(x,y) = \sqrt{|x-y|}$. Is (X,d) a metric space?

1.3 Let $X = \mathbb{R}^n$. Define on X, $d_\infty(x,y) = \max\{|x_k - y_k| : k = 1,\ldots,n\}$ $\forall x = (x_1,\ldots,x_n)$, $y = (y_1,\ldots,y_n)$. Show that (X, d_∞) is a metric space.

1.4 Let d be a metric on X. Define

$$\rho(x,y) = \frac{d(x,y)}{1+d(x,y)} .$$

Show that ρ is a metric on X.

1.5 Two real numbers $p > 1$ and $q > 1$ are called *conjugate exponents,* if

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Show that for all $x, y \in \mathbb{R}_+$ and for conjugate exponents p and q, the following inequality holds.

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Hint: Work with the function $f(z) = \frac{z}{p} + \frac{1}{q} - z^{1/p}$ and then substitute $z = \frac{x^p}{y^q}$.

1.6 (a) Prove Hölder's inequality for finite sums: for conjugate exponents $p > 1$ and $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, $a_1, \ldots, a_n \geq 0$, and $b_1, \ldots, b_n \geq 0$,

$$\sum_{i=1}^{n} a_i b_i \leq \left[\sum_{i=1}^{n} a_i^p \right]^{1/p} \left[\sum_{i=1}^{n} b_i^q \right]^{1/q}.$$

Hint : Apply Problem 1.5 to $x = a_i/A$ and $y = b_i/B$, where

$$A = \left[\sum_{i=1}^{n} a_i^p \right]^{1/p} \text{ and } \quad B = \left[\sum_{i=1}^{n} b_i^q \right]^{1/q}.$$

(b) Generalize Hölder's inequality for infinite sums.

1.7 $a)$ Prove Minkowski's inequality (for finite sums): for $p \geq 1$, $a_1, \ldots, a_n \geq 0$, and $b_1, \ldots, b_n \geq 0$, it holds true that

$$\left[\sum_{i=1}^{n} (a_i + b_i)^p \right]^{1/p} \leq \left[\sum_{i=1}^{n} a_i^p \right]^{1/p} + \left[\sum_{i=1}^{n} b_i^p \right]^{1/p}.$$

(*Hint:* Make use of $(a+b)^p = a(a+b)^{p-1} + b(a+b)^{p-1}$ and then apply Hölder's inequality.)

$b)$ Generalize Minkowski's inequality for infinite sums.

1.8 The *Euclidean metric* or *Euclidean distance* is defined in $\mathbb{R}^n$ by

$$d_e(x, y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}, \tag{P1.8}$$

$$x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n). \tag{P1.8a}$$

[Specifically, if $n = 1$, we have $d(x, y) = \sqrt{(x-y)^2} = |x - y|$.] Show that d_e is indeed a metric (*Hint*: Apply Minkowski's inequality.)

1.9 In Problem 1.8 we defined the Euclidean metric on $\mathbb{R}^n$ by equation (P1.8). This metric can be regarded as

$$d_P(x, y) = \sqrt{\sum_{k=1}^{n} d_k(x_k, y_k)^2}\,, \tag{P1.9}$$

$$x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n), \tag{P1.9a}$$

where $d_k(x_k, y_k)$ is the one-dimensional Euclidean metric on the kth coordinate axis (kth factor space). We can extend this notion and define a metric on the n-times Cartesian product set $Y = Y_1 \times \cdots \times Y_n$ by formulas (P1.9 - P1.9a). This problem asserts that such d_p is indeed a metric on Y. We call this metric the *product metric* and the corresponding metric space (Y, d_P) the product space. In notation, $\times \{(Y_k, d_k) : k = 1, \ldots, n\}$.

Prove the statement: *Let* $(Y_k, d_k),\ k = 1, \ldots, n,$ *be a collection of metric spaces and let* Y *be the Cartesian product of* $Y_1, \ldots, Y_n$. *Then the function* d_P *on* $Y \times Y$ *defined by* (P1.9-P1.9a) *is a metric on* Y.

1.10 Show that the function $\rho(x, y) = \sum_{k=1}^{n} d_k(x_k, y_k)$ is also a metric on $Y = Y_1 \times Y_2 \times \cdots \times Y_n$.

1.11 In Example 1.3 (v), why is $d(x, y) = 0$ if and only if $x = y$?

NEW TERMS:

Fréchet, Maurice, 93
Hausdorff, Felix, 93
Banach, Stefan, 93
metrization, 94
metric, 94
distance, 94
triangle inequality, 94
metric space, 94
pseudo-metric, 94
pseudo-metric space, 94
reducers of a pseudo-metric, 95
$(X|\mathrm{E}_d, \boldsymbol{d})$ (reduced metric space), 95
subspace of a metric space, 95
discrete metric, 95
l^1 space, 96
supremum metric (d_∞), 96
conjugate exponents, 97
Hölder's inequality, 98
Minkowski's inequality, 98
Euclidean metric (d_e), 98
Euclidean distance, 98
product metric, 99
product space, 99

2. THE STRUCTURE OF METRIC SPACES

The structural properties of metric spaces stem from the notion of the open ball, with the aid of which we are be able to introduce open and closed sets, interior, closure, and accumulation points. Open balls, due to a particular metric, generate convergence and continuity, the principles of any analysis, which we explore in this chapter and Chapter 3.

2.1 Definition. Let (X, d) be a metric space and let $x \in X$ and $r > 0$. The subset of X,

$$B(x, r) = \{y \in X : d(x, y) < r\},$$

is called the *open ball centered at x with radius r (with respect to metric d)*. [Whenever we need to emphasize that the ball is with respect to metric d, we write it as $B_d(x, r)$. This notation makes sense if X is endowed with more than one metric.] □

2.2 Examples.

(i) The open ball $B(x, r)$ in Euclidean space $(\mathbb{R}, d_e)$ is the open interval $(x - r,\ x + r)$.

(ii) The open ball $B(x, r)$ in Euclidean space $(\mathbb{R}^2, d_e)$ is the open disc centered at x with radius r in the usual sense.

(iii) Different metric choices on a given carrier give rise to different spaces and, as the result, to different shapes of open balls. Indeed, in metric spaces other than Euclidean, the shape of open balls may be quite surprising to our usual perception of them. Consider, for instance, an open ball $B(x, r)$ in $(\mathbb{R}^2, d)$, where d is the supremum metric defined as in Problem 1.3, for $n = 2$:

$$d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}.$$

It is easy to see that the open ball $B_\infty(x, r)$ is square shaped and that the Euclidean ball $B_e(x, r)$ is inscribed in this square (see Figure 2.1 below).

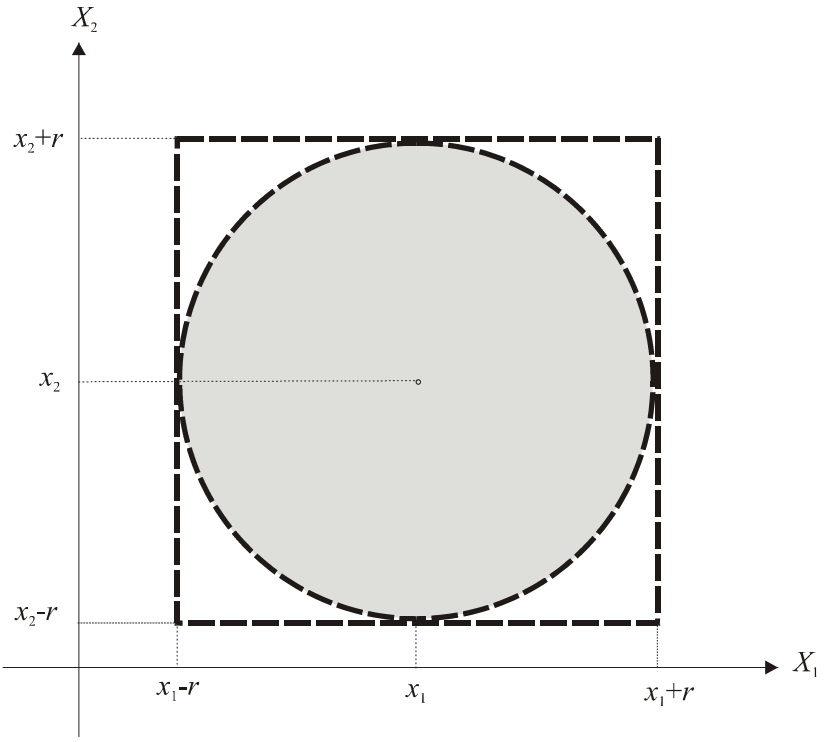


Figure 2.1

(iv) Let (X, d) be a discrete metric space with the metric defined in Example $1.3(i)$. Then, for any $x \in X$, an open ball centered at x is

$$B(x,r) = \begin{cases} \{x\}, & r \leq 1 \\ X, & r > 1. \end{cases}$$

Indeed, because $d(x, y)$ is either 0 or 1, for $r \leq 1$, $B(x, r) = \{y \in X : 0 = d(x, y) < 1\}$, whereas for $r > 1$, $B(x, r) = \{y \in X : d(x, y) = 1\}$.

(v) Let (X, d) be the metric space defined in Example $1.3(iv)$, where $X = C_{[a,b]}$, and

$$d_\infty(x, y) = \sup\{|x(t) - y(t)|: \ t \in [a, b]\}.$$

Then the open ball $B_\infty(x,r)$ has a shape as depicted in Figure 2.2 below.

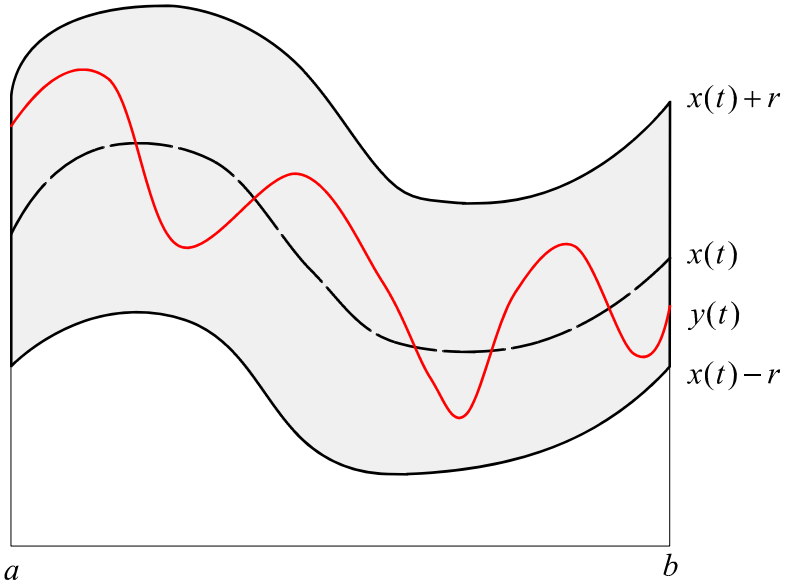


Figure 2.2

□

2.3 Definition. Let (X,d) be a metric space. A subset A of X is called a *d-open set* (or just *open set*) if every point x of A can serve as the center of an open ball inscribed in A, that is, there is an $r>0$ such that $B(x,r)\subseteq A$. □

2.4 Examples.

(i) Every open ball is an open set itself. Indeed, if $x_1\in B(x,r)$ then $r-d(x,x_1)>0$. Take $r_1=r-d(x,x_1)$ and show that $B(x_1,r_1)\subseteq B(x,r)$. For every $z\in B(x_1,r_1)$, by the triangle inequality,

$$d(x,z)\le d(x,x_1)+d(x_1,z)<d(x,x_1)+r_1=r.$$

Thus $z\in B(x,r)$ (see Figure 2.3 below).

(ii) The set $[a, b)$, for $a < b$, in $(\mathbb{R}, d_e)$ is not open, because there is no open ball $B(a, r) \subseteq [a, b)$.

(iii) Any carrier X is obviously open.

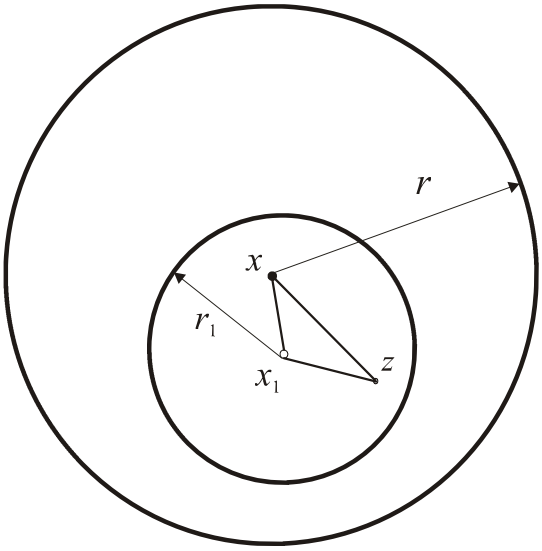


Figure 2.3

(iv) A set A is not open if there is at least one point $x \in A$ such that there is no ball $B(x, r)$ that can be inscribed in A. The empty set does not have any point. Thus it is reasonable to assign it to the class of open sets. In other words, we define Ø as an open set.

(v) In the Euclidean space $(\mathbb{R}, d_e)$, $\mathbb{R}$ is an open set but not an open ball (why?). □

2.5 Theorem. *For every metric space (X, d), the following statements hold true.*

(i) *Arbitrary unions of open sets are open sets.*

(ii) *Finite intersections of open sets are open sets.*

Proof.

(i) Let $\{A_k : k \in I\}$ be an indexed family of open sets in X and let $A = \bigcup_{k\in I} A_k$. If $x \in A$, then there is an i such that $x \in A_i$. Because A_i is open, there is an $r > 0$ such that

$$B(x, r) \subseteq A_i \subseteq \bigcup_{k\in I} A_k .$$

Therefore, A is open.

(ii) Let $A_1, \ldots, A_n$ be open subsets of X and let $A = \bigcap_{k=1}^{n} A_k$. If $x \in A$ then $x \in A_k$, $k = 1, \ldots, n$. It follows that there are $r_1, \ldots, r_n$ such that $B(x, r_k) \subseteq A_k$, $k = 1, \ldots, n$. Let $r = \min\{r_1, \ldots, r_n\}$. Then, obviously, $B(x, r) \neq \emptyset$ and $B(x, r) \subseteq A_k$, $k = 1, \ldots, n$. Thus, $B(x, r) \subseteq A$ and A is open. □

2.6 Remark. The intersection of more than finitely many open sets need not be open. The reason is that because $r = \min\{r_k : k \in I\}$ can be zero. For example, let

$$A_n = (1 - \tfrac{1}{n}, 1 + \tfrac{1}{n}) \text{ in } (\mathbb{R}, d_e).$$

Then $1 \in A_n$, $n = 1, 2, \ldots,$ which implies that $1 \in \bigcap_{n=1}^{\infty} A_n$ and hence $\{1\} = \bigcap_{n=1}^{\infty} A_n$. However, the set $\{1\}$ is not open in $(\mathbb{R}, d_e)$. □

2.7 Example. Let (X, d) be a discrete metric space. Then the power set $\mathcal{P}(X)$ coincides with the set of all open sets. Indeed, in Example 2.2 (iv), we showed that in any discrete metric space, every singleton $\{x\}$ and the carrier X are open balls. In addition, Ø is an open set. Any subset A of X can be represented as the union of all points of A. Thus by Theorem 2.5(i), it follows that A is also open. Specifically, in $\mathbb{R}$ endowed with the discrete metric, all singletons are open, whereas in Euclidean space $(\mathbb{R}, d_e)$ they are not. □

2.8 Definitions.

(i) A point $x \in A \subseteq X$ is called an *interior point* of A if there exists an open ball $B(x, r) \subseteq A$. The set of all interior points of set A is denoted by $\mathring{A}$ or *Int*(A) and called the *interior* of A.

(Clearly, $\mathring{A}$ is the largest open subset of A, which yields that A is open if and only if $A = \mathring{A}$. Indeed, let $C \subseteq A$ be an open set, larger than $\mathring{A}$. Then, there is an $x \in C$ such that $x \notin \mathring{A}$. But this is a contradiction, because x must be an interior point of A.)

(ii) A subset A of X is called *closed* if its complement A^c is open. [Specifically, the carrier X and the empty set Ø are both closed. Any singleton is obviously closed.] □

2.9 Proposition. *Arbitrary intersections or finite unions of closed sets are closed sets*.

Proof. The statements follow from Theorem 2.5 by applying De Morgan's laws. □

2.10 Example. Because the set of all open subsets of a discrete metric space (X, d) coincides with its power set, the set of all closed subsets is also the power set. Particularly, in a discrete metric space all subsets are simultaneously open and closed. □

2.11 Definitions.

(i) A point $x \in X$ is called a *closure point* of $A \subseteq X$ if every open ball centered at x contains at least one element of A. We also say, "if every open ball centered at x *meets* A." The set of all closure points of A is denoted by $\overline{A}$ or by *Cl*(A) and called the *closure* of A. [From Definition 2.8 (i) it follows that $A \subseteq \overline{A}$.]

(ii) There are two types of closures points. Type 1: x is an *isolated closure point of* A if there is a ball $B(x, r)$ that does not contain any points of A, other than x. In other words, $B(x, r) \cap (A \backslash \{x\}) = \text{Ø}$. Clearly, an isolated closure point must belong to A. Type 2: x is an *accumulation point*, if it is not an isolated point. That means, for each $r > 0$, $B(x, r) \cap (A \backslash \{x\}) \neq \text{Ø}$. Unlike an isolated closure point, an accumulation point need not belong to A.

The set of all accumulation points of A is called a *derived set*, in notation A'. From the above definition it follows that $A' \subseteq \overline{A}$.

(iii) A point $x \in X$ is called a *boundary point* of A if every open ball at x contains points from A and from A^c. The set of all boundary points of A is called the *boundary* of A and is denoted by ∂A. □

2.12 Examples.

(i) In $(\mathbb{R}, d_e)$, let $A = (0,2) \cup \{5\}$. $\{5\}$ is an isolated closure point, because $B(5,r)$ contains only $\{5\}$ for all $r \leq 3$. Thus, $\overline{A} = [0,2] \cup \{5\}$, and $A' = [0,2]$.

(ii) Note that $\partial A = \partial A^c = \overline{A} \cap \overline{A^c}$. (See Problem 2.1.) Therefore, ∂A (as follows from Proposition 2.13 below) is a closed set.

(iii) In the next section we learn that the closure $\overline{\mathbb{Q}}$ of all rational numbers in $(\mathbb{R}, d_e)$ is entire $\mathbb{R}$. So is the closure $\overline{\mathbb{Q}^c}$ of irrational numbers. Consequently, by (ii), $\partial\mathbb{Q} = \mathbb{R}$, which is quite surprising. [Note that $\partial\mathbb{Q} = \mathbb{R}$ also directly follows from Definition 2.11(iii).]

(iv) The boundaries of X and Ø are the empty sets.

(v) Let $A = [0,1) \cup \{2\}$. Then, $\mathring{A} = (0,1)$, $\overline{A} = [0,1] \cup \{2\}$, $A' = [0,1]$, $\partial A = \{0,1,2\}$ (because $A^c = (-\infty, 0) \cup [1,2) \cup (2,\infty)$, $\overline{A^c} = (-\infty, 0] \cup [1,\infty)$, and $\overline{A} \cap \overline{A^c} = \{0,1,2\}$).

(vi) Let $A = \{1, \frac{1}{2}, \frac{1}{3}, \dots\} \subseteq (\mathbb{R}, d_e)$. Because 0 is the limit of the sequence $\{\frac{1}{n}\}$ (in terms of Euclidean distance), it is also an accumulation point of A. Any open ball at 0 contains at least one point of A. This is the only accumulation point of A. By the way, A is not closed, for 0 is a closure point of A. So we have $A' = \{0\}$, and $\overline{A} = A \cup \{0\}$. □

2.13 Proposition. *For any subset A of X, $\overline{A}$ is a closed set and furthermore, $\overline{A}$ is the smallest closed superset of A.*

Proof.

(i) We show first that $\overline{A}$ is a closed set, that is, $(Cl(A))^c$ is open. Let $x \in (Cl(A))^c$. Then there exists an open ball $B(x,r)$ such that $B(x,r) \cap A = \text{Ø}$ (because, otherwise, x would belong to $\overline{A}$ by the definition). However, we have not proved yet that $B(x,r) \cap \overline{A} = \text{Ø}$, which would immediately imply that $(Cl(A))^c$ is open. Now we show that no point of $B(x,r)$ is a closure point of A. Take an arbitrary point $t \in B(x,r)$. Because $B(x,r)$ is an open set, there is an open ball $B(t,r_t) \subseteq B(x,r)$ also disjoint from A. By the

definition of a closure point, this means that $t \notin \overline{A}$. In as much as t was an arbitrary point of $B(x,r)$, $B(x,r) \subseteq (Cl(A))^c$.

(ii) Now we show that the closure of A is the smallest closed set containing A. Let B be an arbitrary closed set such that $A \subseteq B$. We prove that $B^c \subseteq (\overline{A})^c$. Because B^c is open, for each $x \in B^c$, there is an open ball $B(x,r) \subseteq B^c$. This implies that $B(x,r) \cap B = \emptyset$ and that $B(x,r) \cap A = \emptyset$.

Thus $x \notin \overline{A}$ (by the definition of a closure point), which is equivalent to $x \in (Cl(A))^c$. Therefore, we have proved that $x \in B^c$ yields that $x \in (Cl(A))^c$, that is, $B^c \subseteq (Cl(A))^c$. The latter is obviously equivalent to $\overline{A} \subseteq B$. □

2.14 Corollary. *A set A is closed if and only if $A = \overline{A}$.*

(See Problem 2.15.)

2.15 Remark. Consider the set $C = C(x,r) = \{y \in X : d(x,y) \leq r\}$. It can be easily shown that C is a closed set. (See Problem 2.4.) Such a C is called a *closed ball centered at x with radius r*. Evidently, $B(x,r) \subseteq C(x,r)$ implies that the closure of an open ball is a subset of C [i.e., $\overline{B(x,r)} \subseteq C(x,r)$], because $\overline{B}$ is the smallest closed set containing B. However, $C(x,r)$ need not coincide with the closure of the corresponding open ball $B(x,r)$. For instance, let (X,d) be a discrete metric space, where any open ball is both a closed and open set: $B(x,r) = \overline{B(x,r)}$. Because

$$C(x,r) = \begin{cases} \{x\}, & r < 1 \\ X, & r \geq 1, \end{cases}$$

we have $B(x,r) = C(x,r) = X$ for $r > 1$ or $B(x,r) = C(x,r) = \{x\}$ for $r < 1$. For $r = 1$, $\overline{B(x,r)} = \{x\} \subset C(x,r) = X$, unless X is a singleton. □

2.16 Example. The set of all rational numbers $\mathbb{Q}$ is neither open nor closed. Indeed, it is known that each irrational point x is a limit of a sequence of rational points $\{x_n\}$. Therefore, if x is irrational, there is no open ball $B(x,r)$ that does not contain rational points. This implies that $\mathbb{Q}^c$ is not open either, or equivalently, $\mathbb{Q}$ is not closed. On the other hand, $\mathbb{Q}$ cannot be open, because otherwise,

every rational point q could be the center of an open ball (interval) containing just rational numbers. This is absurd, in as much as any interval is a continuum. Therefore, the set of all rational numbers is neither open nor closed. It also follows that the set of all irrational numbers is neither open nor closed. □

In Problem 1.9, we introduced a product metric. We wonder how open sets can look like in a product metric space. A remarkable property of such a metric is given by the following theorem to be proved in Chapter 3 for the general topological product spaces.

2.17 Theorem. *Let* $\{(Y_k, d_k) : k = 1, \ldots, n\}$ *be a finite family of metric spaces and let* $(Y, d) = \times \{(Y_k, d_k) : k = 1, \ldots, n\}$ *be the product space. Then* $O \subseteq (Y, d)$ *is open if and only if* O *is the union of sets of the form* $\prod_{i=1}^{n} O_i$, where each O_i is open in (Y_i, d_i).

PROBLEMS

2.1 Show that $\partial A = \partial A^c = \overline{A} \cap \overline{A^c}$.

2.2 Is it true that $A \subseteq B$ yields

(a) $\overline{A} \subseteq B$?

(b) $\overline{A} \subseteq \overline{B}$?

2.3 Show that $[\overline{A^c}]^c \subseteq \overline{A}$.

2.4 Prove that a closed ball $C(x, r)$ is a closed set.

2.5 If $x \in \partial A$, must x be an accumulation point?

2.6 Show that $\overline{A} = A \cup A'$.

2.7 Let $A \subseteq (X, d)$, where X is an infinite set. Show that if x is an accumulation point of A, then every open set containing x contains infinitely many points of A.

2.8 Give an example of a continuum closed set that does not have any accumulation point.

2.9 Describe the shape of open balls in the metric space (X, d) introduced in Example 1.3(ii).

2.10 Show that set $[1, \infty)$ is closed in the metric space of Example 1.3(ii).

2.11 Prove that $\overline{A} = \mathring{A} + \partial A$.

2.12 Show that $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

2.13 Show that $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$.

2.14 Show that $\overline{A \cap B} \neq \overline{A} \cap \overline{B}$.

2.15 Prove Corollary 2.14: *A set A is closed if and only if $A = \overline{A}$.*

2.16 Show that in any discrete metric space, the boundary of any subset is empty.

2.17 Define $S(x, r) = \{y \in X : d(x, y) = r\}$ and call it a *sphere centered at x with radius r $(r > 0)$*. In Euclidean space $(\mathbb{R}^n, d_e)$, obviously, $S(x, r) = \partial B(x, r)$. Is this true for metric spaces? If it is not, then is there at least a relationship between $S(x, r)$ and $\partial B(x, r)$ common for all metric spaces? Are there any exceptional cases? Explain.

2.18 Show that set $(0, 1]$ is closed in the metric space of Example 1.3 (ii).

2.19 Let O be an open set in $(\mathbb{R}^n, d_e)$. Show that for each $x \in O$, there is an open ball $B(q, r) \ni x$, such that $q \in \mathbb{Q}^n$ (the set of all vectors in $\mathbb{R}^n$ with rational coordinates) and $r \in \mathbb{Q}$.

NEW TERMS:

open ball $B(x, r)$, 101
radius of an open ball, 101
supremum metric,101, 102
open ball with respect to the Euclidean metric, 102
open ball with respect to the supremum metric, 102, 103
open (d-open) set, 103
interior point, 105
interior $\mathring{A}$ or *Int*(A) of set A, 105
closed set, 106
closure point, 106
closure $\overline{A}$ or *Cl*(A) of set A, 106
isolated closure point, 106
accumulation point, 106
derived set (A'), 107
boundary point, 107
boundary ∂A of set A, 107
closed ball $C(x, r)$, 108
sphere $S(x, r)$, 110

3. CONVERGENCE IN METRIC SPACES

This section introduces the reader to one of the central notions in the analysis of metric spaces: convergence. Among different things, we discuss the relation between limit and closure points.

3.1 Definitions.

(i) A function $[\mathbb{N}, X, f]$ is called a *sequence*; its most commonly used notation is $\{x_n\} = f$, with $x_n = f(n)$. Let $\{x_n\} \subseteq (X, d)$ be a sequence and let $\varepsilon > 0$ and N be a positive integer. A subsequence $T_N = \{x_N, x_{N+1}, \dots\}$ is called an $N(\varepsilon)$-*tail* of $\{x_n\}$ if for any $x_m, x_n \in T_N$, $d(x_m, x_n) < \varepsilon$. The sequence $\{x_n\}$ is called a *Cauchy sequence*, in notation

$$\lim_{m,n\to\infty} d(x_m, x_n) = 0,$$

if for every $\varepsilon > 0$, there is an $N(\varepsilon)$-tail of $\{x_n\}$.

(ii) Let $x \in X$. Any $N(\varepsilon)$-tail, T_N, is called an $N(x, \varepsilon)$-*tail of* $\{x_n\}$ if $T_N \subseteq B(x, \varepsilon)$. The sequence $\{x_n\}$ is said to *converge* to a point $x \in X$ if for every $\varepsilon > 0$, there is an $N(x, \varepsilon)$-tail. In notation,

$$\lim_{n\to\infty} d(x_n, x) = 0$$

(also $d\text{-}\lim_{n\to\infty} x_n = x$ or just $x_n \to x$). x is called a *limit point* of the sequence $\{x_n\}$. A sequence is *convergent* if it is convergent to at least one limit point that belongs to X.

(iii) A point x is said to be a *sequential limit point of a set* A if there is a sequence $\{x_n\} \subseteq A$ convergent to x.

(iv) A set A in a metric space (X, d) is called *complete* if every Cauchy sequence in A is convergent in A. Consequently, the metric space (X, d) is complete if so is X.

(v) A sequence $\{x_n\}$ is called *bounded* if for every n, $d(x_1, x_n) \leq M$, where M is a positive real number. □

3.2 Remark. A sequence in a metric space can have at most one limit point. Indeed, let x, y be limits of a sequence $\{x_n\} \subseteq (X, d)$

and let $\varepsilon > 0$ be arbitrary. Then, there are $N_1(x, \varepsilon/2)$ and $N_2(y, \varepsilon/2)$ tails such that

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon,$$

for all $n \geq N = \max\{N_1, N_2\}$ [i.e., $d(x, y)$ can be made arbitrarily small]. Thus, $x = y$. □

3.3 Theorem. *Let $A \subseteq (X, d)$. Then a point x is a closure point of a set A if and only if x is a sequential limit point of A (i.e., there is a sequence $\{x_n\} \subseteq A$ such that $x_n \to x$).*

Proof.

(i) Let x be a closure point of A. If $x \in A$ then the proof becomes trivial (if we set $x_n = x,\ n = 1, 2, \dots$). Let $x \in \overline{A} \backslash A$. By the definition of a closure point, every open ball $B(x, r)$ meets A. Thus for every n, there is a point, $x_n \in A \cap B(x, \frac{1}{n})$, so that $d(x, x_n) < \frac{1}{n}$. Therefore, $\{x_n\}$ is a pledged sequence convergent to x.

(ii) Let $\{x_n\} \subseteq A$ such that $\lim\limits_{n \to \infty} x_n = x$. We prove that $x \in \overline{A}$. The convergence implies that for every $\varepsilon > 0$, there is a $T_N = \{x_N, x_{N+1}, \dots\}$ such that $T_N \subseteq B(x, \varepsilon)$ and $T_N \subseteq A$. Thus, $\forall \varepsilon > 0,\ B(x, \varepsilon) \cap A \neq \emptyset$, which yields that $x \in \overline{A}$.

Part (i) implies that if A is an infinite set and $x \in A' \backslash A \neq \emptyset$, i.e., x is an accumulation point of A, then there is a sequence $\{x_n\}$ with all distinct terms such that $x_n \to x$.) □

3.4 Corollary. *A subset A of a metric space (X, d) is closed if and only if it contains all of its sequential limit points.*

Proof.

(i) Let A be closed and let $\{x_n\} \subseteq A$ be a sequence convergent to a point x. Then, by Theorem 3.3, x is a sequential limit point and a closure point of A at the same time:

$$\lim_{n \to \infty} x_n = x \in \overline{A}.$$

Since A is closed, $A = \overline{A}$ and $x \in A$. Thus, A contains all of its sequential limit points.

(ii) Suppose A contains all of its sequential limit points and let $x \in \overline{A}$. Then, by Theorem 3.3, there is a sequence $\{x_n\} \subseteq A$ such that $\lim_{n \to \infty} x_n = x$. By our assumption, x also belongs to A or, equivalently, $\overline{A} \subseteq A$ implying that $A = \overline{A}$ and hence A is closed. □

3.5 Definitions.

(i) A subset $A \subseteq (X, d)$ is called *dense* in X, if $\overline{A} = X$. (By Theorem 3.3, A is dense in X if and only if the set of all limit points of A coincides with X, or, in other words, if and only if for every $x \in X$, there exists a sequence $\{x_n\} \subseteq A$ such that $x_n \to x$.)

(ii) A set $A \subseteq (X, d)$ is called *nowhere dense* if its closure has the empty set as its interior, that is, if *Int*$(Cl(A)) = \emptyset$. □

3.6 Examples.

(i) Because each irrational number can be represented as the limit of a sequence of rational numbers, $\mathbb{Q}$ is dense in $\mathbb{R}$ (in terms of the Euclidean metric).

(ii) Let $A = \{1, 5, 10\} \subseteq (\mathbb{R}, d_e)$. Then A is nowhere dense.

(iii) $\{\frac{1}{n} : n = 1, 2, \dots\}$ is nowhere dense in $(\mathbb{R}, d_e)$.

PROBLEMS

3.1 Show that every convergent sequence is a Cauchy sequence. Give an example when the converse is not true.

3.2 Prove that a set $A \subseteq (X, d)$ is nowhere dense in X if and only if the complement of its closure is dense in X.

3.3 Assuming that $(\mathbb{R}, d_e)$ is complete (a known fact from calculus), prove that $(\mathbb{R}^n, d_e)$ is also complete.

3.4 Show that any Cauchy sequence is bounded.

3.5 Show that in a discrete metric space any convergent sequence has at most finitely many distinct terms.

3.6 Show that any discrete metric space is complete.

3.7 Show that if $\{x_n\} \subseteq (X, d)$ is a Cauchy sequence and $\{x_{n_k}\}$ is a subsequence convergent to a point $a \in X$, then $x_n \to a$.

3.8 Show that in $(\mathbb{R}^n, d_e)$, $\overline{B(x,r)} = C(x,r)$, where $C(x,r)$ is a closed ball of Remark 2.15.

NEW TERMS:

sequence, 112
$N(x,\varepsilon)$-tail, 112
Cauchy sequence, 112
convergent sequence, 112
limit point of a sequence, 112
sequential limit point of a set, 112
complete metric space, 112
bounded sequence, 112
dense set, 114
nowhere dense set, 114

4. CONTINUOUS MAPPINGS IN METRIC SPACES

4.1 Definition. Let (X, d) and (Y, ρ) be two metric spaces. A function $f : (X, d) \to (Y, \rho)$ is called *continuous at a point* $x_0 \in X$ if for each $\varepsilon > 0$, there is $\delta > 0$ such that $\rho(f(x), f(x_0)) < \varepsilon$ for all x with $d(x, x_0) < \delta$. Function f is called *continuous on* X or simply *continuous* if f is continuous at every point of X. □

4.2 Remark. (The anatomy of continuity.) Since for any function, $x_0 \in f^*\Big(\{f(x_0)\}\Big)$, we always have $x_0 \in f^*(B_\rho(f(x_0), \varepsilon))$ [the inverse image of an open ball in Y centered at $f(x_0)$], the key question is whether x_0 is an interior point of $f^*(B_\rho(f(x_0), \varepsilon))$ for any $\varepsilon > 0$. We show that f is continuous at x_0 if and only if the inverse image under f^* of any open ball centered at $f(x_0)$ contains x_0 as an interior point. (See Figure 4.1.)

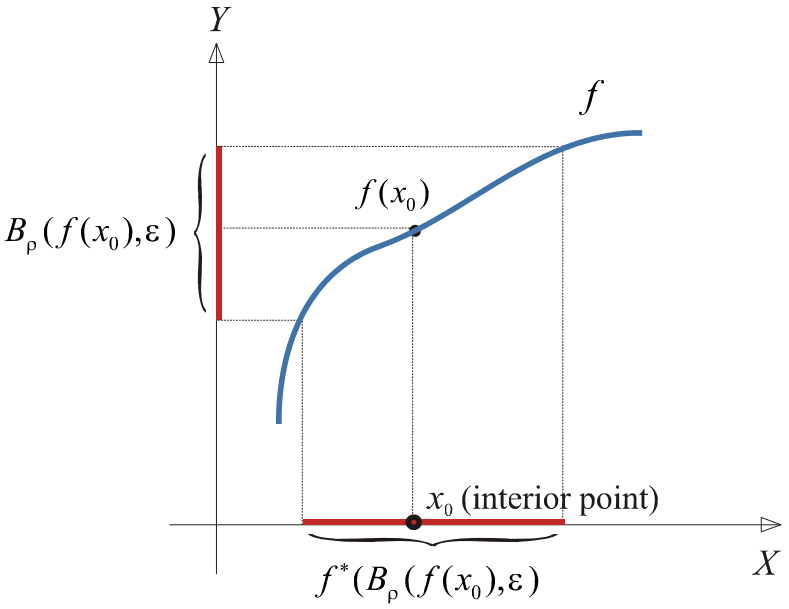


Figure 4.1

If x_0 is an interior point of $f^*(B_\rho(f(x_0),\varepsilon))$, as conjectured, there must be an open ball $B_d(x_0,\delta) \subseteq f^*(B_\rho(f(x_0),\varepsilon))$. It implies that: 1) such a positive δ exists, and 2) the image of $B_d(x_0,\delta)$ under f_* is a subset of $B_\rho(f(x_0),\varepsilon)$ (see arguments in Theorem 4.3 below), which guarantees that $\rho(f(x),f(x_0)) < \varepsilon$ for all x with $d(x,x_0) < \delta$.

Now, if f is not continuous at x_0, as depicted in Figure 4.2, x_0 need not be an interior point of $f^*(B_\rho(f(x_0),\varepsilon))$. If this is the case, no ball $B_d(x_0,\delta)$ can be inscribed in $f^*(B_\rho(f(x_0),\varepsilon))$ and consequently, no positive δ exists to warrant that $\rho(f(x),f(x_0))$ is less than ε for all x with $d(x,x_0) < \delta$.

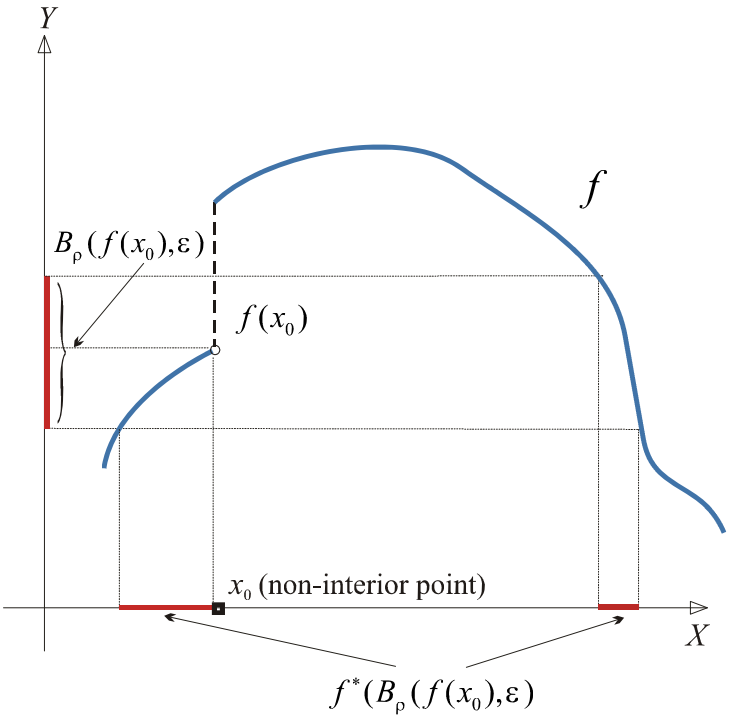


Figure 4.2

On the other hand, the reader should not jump to a quick conclusion that Figure 4.2 (which serves for a mere illustration) necessarily represents a discontinuous function. For instance, if X is endowed

with the discrete metric, then x_0 is an interior point of the set $f^*(B_\rho(f(x_0), \varepsilon))$, because $\{x_0\}$ is an open ball with any radius $r \leq 1$. Furthermore, $f^*(B_\rho(f(x_0), \varepsilon))$ is an open set in X for all $\varepsilon > 0$. □

The following theorem elaborates the above discussion and also generalizes the principle of continuity.

4.3 Theorem. *A function $f : (X, d) \to (Y, \rho)$ is continuous if and only if the inverse image of any open set in (Y, ρ) under f is open in (X, d).*

Proof.

1) According to the blueprint in Remark 4.2, we begin the proof by showing the validity of the following assertion.

> *f is continuous at x_0 if and only if x_0 is an interior point of the inverse image under f^* of any open ball $B_\rho(f(x_0), \varepsilon)$.*

Let x_0 be an interior point of $f^*(B_\rho(f(x_0), \varepsilon))$. Then, there is an open ball

$$B_d(x_0, \delta) \subseteq f^*(B_\rho(f(x_0), \varepsilon)),$$

and because inclusions are preserved under f_*, by Problems 3.6(a) and 2.6 of Chapter 1,

$$f_*(B_d(x_0, \delta)) \subseteq f_*(f^*(B_\rho(f(x_0), \varepsilon))) \subseteq B_\rho(f(x_0), \varepsilon),$$

which yields continuity of f at x_0, as per Remark 4.2.

Now, let f be continuous at x_0. Then, the inclusion $f_*(B_d(x_0, \delta)) \subseteq B_\rho(f(x_0), \varepsilon)$ holds, which, along with Problem 2.5, Chapter 1, leads to the following sequence of inclusions.

$$B_d(x_0, \delta) \subseteq f^*(f_*(B_d(x_0, \delta))) \subseteq f^*(B_\rho(f(x_0), \varepsilon)).$$

Because x_0 is the center of $B_d(x_0, \delta)$, it is an interior point of this ball and, due to the last inclusion, an interior point of the set $f^*(B_\rho(f(x_0), \varepsilon))$.

2) Suppose f is continuous on X. We show that for each open set $O \subseteq Y$, $f^*(O)$ is open in (X, d). Pick a point $x_0 \in f^*(O)$. Then, $f(x_0) \in f_*(f^*(O)) \subseteq O$ and, because O is open, $f(x_0)$ is its interior point. Thus, O is a superset of the open ball $B_\rho(f(x_0), \varepsilon)$, for some ε, and consequently,

$$f^*(B_\rho(f(x_0), \varepsilon)) \subseteq f^*(O). \tag{4.3}$$

Because f is continuous at x_0, by assertion 1), x_0 must be an interior point of $f^*(B_\rho(f(x_0), \varepsilon))$, and, by (4.3), an interior point of $f^*(O)$. Thus, $f^*(O)$ is open.

3) Let $f^*(O)$ be open in (X, d) for every open subset O of Y. Take $x_0 \in X$ and construct an open ball $B_\rho(f(x_0), \varepsilon)$. By our assumption, the set $f^*(B_\rho(f(x_0), \varepsilon))$ is open in (X, d). Because $f(x_0) \in B_\rho(f(x_0), \varepsilon)$, we have that

$$x_0 \in f^*(\{f(x_0)\}) \subseteq f^*(B_\rho(f(x_0), \varepsilon))$$

and, therefore, $x_0 \in f^*(B_\rho(f(x_0), \varepsilon))$ and it is an interior point of $f^*(B_\rho(f(x_0), \varepsilon))$. By 1), f must then be continuous at x_0. □

There is yet another useful criterion of continuity, known as Heine's criterion (after the German mathematician Heinrich Eduard Heine, 1821 – 1861).

4.4 Theorem (H.E. Heine). *A function $f\colon (X, d) \to (Y, \rho)$ is continuous at $x \in X$ if and only if for every sequence $\{x_n\}$, d-convergent to x, its image sequence $\{f(x_n)\}$ is ρ-convergent to $f(x)$.*

We prove a version of this theorem for a more general case in Chapter 3 (Theorems 4.9 and 4.10).

4.5 Definition. Let (X, d) be a metric space and $\tau(d)$ be the collection of all open subsets of X with respect to metric d. Then $\tau(d)$ (or just τ) is said to be the *topology on X generated by d.* □

Theorem 4.3 can now be rephrased as follows.

4.6 Theorem. *Let $f\colon (X, d) \to (Y, \rho)$ be a function and let $\tau(d)$ and $\tau(\rho)$ be the topologies generated by metrics d and ρ,*

respectively. Then f *is continuous on* X *if and only if* $f^{**}(\tau(\rho)) \subseteq \tau(d)$ [*i.e.*, $\forall O \in \tau(\rho),\ f^*(O) \in \tau(d)$]. □

4.7 Example. Let $f\colon (\mathbb{R}, d) \to (\mathbb{R}, d_e)$ be the *Dirichlet function* defined as $f = \mathbf{1}_{\mathbb{Q}}$, where $\mathbb{Q}$ is the set of rational numbers. If $d = d_e$ is the Euclidean metric then f is discontinuous at every point. If d is the discrete metric, by Theorem 4.3, f is continuous on $\mathbb{R}$, because the inverse image of any open set in $(\mathbb{R}, d_e)$ under f is clearly an element of the power set coinciding with the *discrete topology* generated by the discrete metric (see Example 2.7). □

We are further interested in the conditions under which two different metrics on X generate one and the same topology. This property of metrics satisfies an equivalence relation on the set of all topologies on X and hence is regarded as equivalence of metrics.

4.8 Definition. Two metrics d_1 and d_2 on X are called *equivalent* if $\tau(d_1) = \tau(d_2)$ (in notation $d_1 \approx d_2$). □

According to Definition 4.8, all metrics sharing one and the same topology τ on X form one equivalence class, say $[d]_\tau$. Therefore, if $\mathfrak{D}_X$ is the set of all metrics on X, then $\mathfrak{D}_X|\mathrm{E}_\tau$ is the quotient set of $\mathfrak{D}_X$ modulo E_τ.

4.9 Remark. Let (X, d_1) and (X, d_2) be two metric spaces and let $f\colon (X, d_1) \to (X, d_2)$ be the identity map $(f(x) = x,\ x \in X)$. If d_1 and d_2 are equivalent [i.e., $\tau(d_1) = \tau(d_2)$], then for every open set O in (X, d_2), $f^*(O) = O \in \tau(d_1)$. Thus f is continuous on X. According to Heine's Theorem 4.4, this is equivalent to the statement that

$$\lim_{n \to \infty} d_1(x_n, x) = 0$$

implying that

$$\lim_{n \to \infty} d_2(f(x_n), f(x)) = \lim_{n \to \infty} d_2(x_n, x) = 0.$$

In a nutshell, assuming that

(i) $\tau(d_1) = \tau(d_2)$

we showed that

$$(ii) \quad \lim_{n\to\infty} d_1(x_n, x) = 0 \;\Leftrightarrow\; \lim_{n\to\infty} d_2(x_n, x) = 0.$$

By Heine's theorem, it follows that the converse is also true: statement (ii) implies statement (i). Hence, we may call two metrics $\tau(d_1)$ and $\tau(d_2)$ on X equivalent if (i) or (ii) holds.

In short, two metrics d_1 and d_2 on X are equivalent if and only if any sequence $\{x_n\} \subseteq X$ convergent in d_1 is convergent in d_2 and vice versa. □

From Theorem 4.3, it also follows that the identity map in Remark 4.9 is continuous "under equivalent metrics." However, an identity map need not be continuous under d_1 and d_2 if they are not equivalent.

4.10 Definitions.

(i) Let A be a subset in a metric space (X, d). The real number or infinity

$$d(A) = \sup\{d(x, y) : x, y \in A\}$$

is called the *diameter* of A. The set A is called *d-bounded* or just *bounded* if $d(A) < \infty$. Particularly, the metric space (X, d) (or just metric d) is called *bounded* if so is X. Set A is said to be *unbounded* if $d(A) = \infty$.

(ii) A subset A in a metric space (X, d) is called *totally bounded* if for every $\varepsilon > 0$, A can be covered by finitely many ε-balls (i.e., balls with common radius ε). □

4.11 Example. According to Problem 1.4, the function

$$\rho(x, y) = \frac{d(x,y)}{1 + d(x,y)}$$

defined on a metric space (X, d) is a metric on X. Obviously

$$\lim_{n\to\infty} d(x_n, x) = 0$$

if and only if $\lim_{n\to\infty} \rho(x_n, x) = 0$ (due to $d = \frac{\rho}{1-\rho}$). Therefore, d and ρ are equivalent. Observe that ρ is clearly bounded whereas d is arbitrary. □

We close this section with a short discussion on uniform continuity. This concept is developed further in Section 6 and Chapter 3. Uniform continuity is an important notion in analysis. Real-valued uniformly continuous functions can be approximated by step functions and this is a core issue when proving the Riemann integrability of continuous functions. The generalization of uniform continuity for metric spaces is useful in more abstract settings and is relatively straightforward.

4.12 Definition. A function $f\colon (X, d) \to (Y, \rho)$ is called *uniformly continuous* on X if for every $\varepsilon > 0$, there is a positive real number δ such that $d(x, y) < \delta$ implies that $\rho(f(x), f(y)) < \varepsilon$, for every $x, y \in X$. □

In the case of usual continuity, a delta depends upon a particular point $x \in X$, where the continuity holds, so that a common delta, good for all points $x \in X$, need not exist. Unlike continuity, uniform continuity guarantees the existence of such positive δ (for every fixed ε) and for all points of X simultaneously. Clearly, uniform continuity implies continuity.

Uniform continuity can also be defined on some subset A of X, so that in Definition 4.12, X will be replaced with A.

4.13 Examples.

(i) Consider $f : (\mathbb{R}, d_e) \to (\mathbb{R}, d_e)$ such that $f(x) = x^2$. Then

$$|x_0 - x| < \delta$$

implies that

$$|x + x_0| = |x - x_0 + 2x_0|$$

$$\leq |x - x_0| + 2|x_0| < \delta + 2|x_0|$$

and

$$|f(x) - f(x_0)| = |x^2 - x_0^2| = |x - x_0| \cdot |x + x_0|$$

$$< \delta \cdot (\delta + 2|x_0|).$$

Take $\delta \cdot (\delta + 2|x_0|)$ as ε. Then δ can be found explicitly as a function of ε such that

$$\delta = \sqrt{x_0^2 + \varepsilon} - |x_0|.$$

Therefore, function $x \mapsto x^2$ is d_e-continuous at every point $x_0 \in \mathbb{R}$. However, x^2 is not uniformly continuous on $\mathbb{R}$, because δ depends upon x_0 as well. Specifically, $\delta \to 0$ when $x_0 \to \infty$. Consequently, we cannot find a $\delta > 0$ good for all x_0.

(ii) Let $f(x) = x^2$ be given as

$$f : ([0, 3], d_e) \to ([0, 9], d_e).$$

From the last inequality above we derive

$$| f(x) - f(x_0) | < \delta(\delta + 2|x_0|) \leq \delta(\delta + 6).$$

The latter inequality implies that $\delta = \sqrt{9 + \varepsilon} - 3$, where $\varepsilon = \delta(\delta + 6)$. Thus $d_e(f(x), f(x_0)) < \varepsilon$ whenever $d_e(x, x_0) < \delta = \sqrt{9 + \varepsilon} - 3$. Because δ is independent of x_0, $f(x)$ is uniformly continuous. Observe that f has been given on a closed and bounded interval which provides the uniform continuity. However, in this case, f would also be uniformly continuous even if f were defined on any bounded but not necessarily closed interval, for instance $(0, 3)$ (why?).

(iii) A continuous function can be uniformly continuous over unbounded sets as, for example, functions $f(x) = \frac{1}{x}$, $x \in [1, \infty)$, and $f(x) = \sin x$, $x \in \mathbb{R}$.

(iv) If (X, d) and (Y, ρ) are two metric spaces, an *isometry* from X to Y is any bijective map $[X, Y, f]$ such that

$$d(x_1, x_2) = \rho(f(x_1), f(x_2)), \text{ for all } x_1, x_2 \in X.$$

Obviously, such an f is uniformly continuous on X.

(v) A function f from (X,d) to (Y,ρ) is said to *satisfy the Lipschitz condition on* X if there is a nonnegative constant C such that for each $x,y\in X, \rho(f(x),f(y))\le Cd(x,y)$. The smallest such C is called the *Lipschitz constant of* f. The function f is uniformly continuous on X if f satisfies the Lipschitz condition. (See Problem 4.17.) □

There is an analytical result, known as the Heine-Borel theorem, stating that any continuous function defined on a closed and bounded set in any Euclidean metric space is also uniformly continuous. The general form of this result will be discussed in Section 6 (Theorem 6.13).

4.14 Definition. A function f from (X,d) to (Y,ρ) is said to be *Cauchy continuous* if the image under f_* of every d-Cauchy sequence $\{x_n\}$ [i.e., $f_*(\{x_n\})$] is a ρ-Cauchy sequence. □

4.15 Example. Notice that continuous functions do not map Cauchy sequences to Cauchy sequences, that is, continuous functions need not be Cauchy continuous. Let us introduce the function $[(-\frac{\pi}{2},\frac{\pi}{2}),\mathbb{R},\text{rtan}]$, where $\text{rtan} := \text{Res}_{(-\frac{\pi}{2},\frac{\pi}{2})}\tan$, and set $\arctan := \text{rtan}^{-1}$. Consider the sequence

$$x_n := \arctan(n),\ n=1,2,\dots\ .$$

Then, although $\{x_n\}$ is a Cauchy sequence in $(\mathbb{R},d_e)$, $\{\text{rtan}(x_n)\}$ is not. (See Problem 4.14.) □

Notice that continuous functions do not map Cauchy sequences to Cauchy sequences as we learned it from Example 4.15. However, uniformly continuous functions do it.

The statement below is an easy exercise. (See Problem 4.15.)

4.16 Proposition. *If f from (X,d) to (Y,ρ) is uniformly continuous, then it is also Cauchy continuous*. □

4.17 Remark. Suppose $f:(X,d)\to(Y,\rho)$ is a map. What can go wrong with f if it is not uniformly continuous? Suppose f is

continuous, but not uniformly continuous. Thus, given an $\varepsilon > 0$ and an x_1, there is a $\delta_1 := \delta(x_1, \varepsilon)$ such that $f_*(B_d(x_1, \delta_1)) \subseteq B_\rho(f(x_1), \varepsilon)$. Using the same ε, we can find for another point x_2 a $\delta_2 := \delta(x_2, \varepsilon)$ such that $f_*(B_d(x_2, \delta_2)) \subseteq B_\rho(f(x_2), \varepsilon)$. If we continue with this process point after point, it may happen that along some sequential path $\{x_n\}$, it is getting increasingly harder to "squeeze" $f_*(B_d(x_n, \delta_n))$ into $B_\rho(f(x_n), \varepsilon)$ while δ_n becomes smaller and smaller, as we can see in Figure 4.3 below.

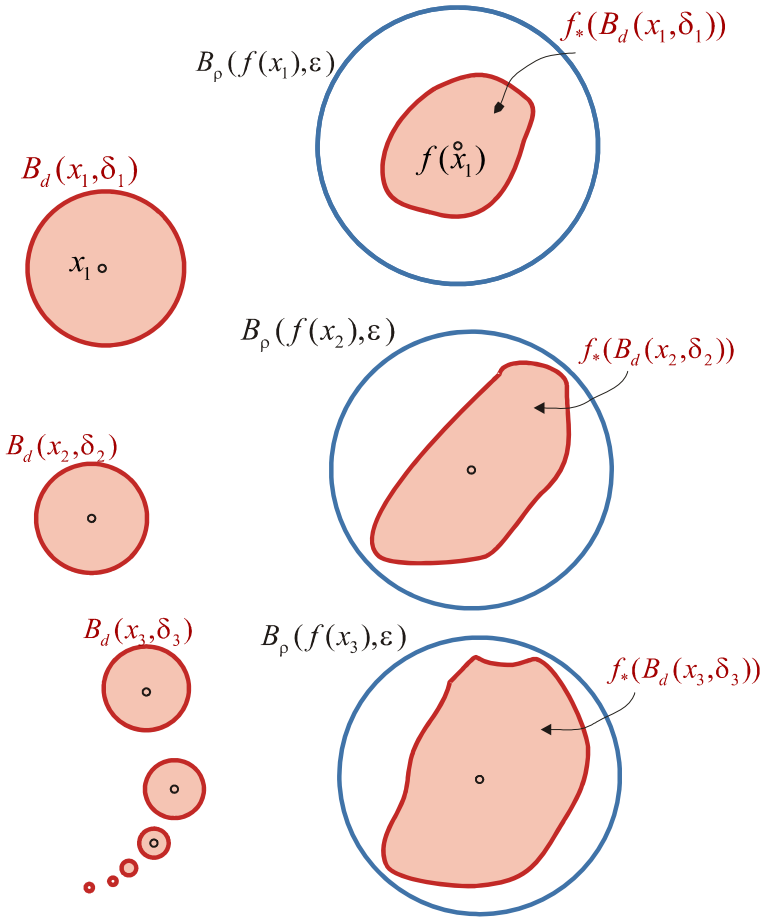


Figure 4.3

So, along the path $\{x_n\}$, $\delta_n \downarrow 0$. The sequence $\{x_n\}$ was brought for an illustration. In general, if f is not uniformly continuous and X is an infinite set, then there is no positive $\delta(\varepsilon)$ good for all $x \in X$ such that $f_*(B_d(x, \delta(\varepsilon))) \subseteq B_\rho(f(x), \varepsilon)$. However, if such a positive $\delta(\varepsilon)$ does exist, f is uniformly continuous.

There is another way to say this. Suppose f is continuous on X and let $\varepsilon > 0$ be chosen. Then for each $x \in X$, there is a $\delta_x = \delta_x(\varepsilon) > 0$ such that $f_*(B_d(x, \delta_x)) \subseteq B_\rho(f(x), \varepsilon)$. Let $\Delta := \{\delta_x(\varepsilon) : x \in X\}$. If f is not uniformly continuous, then for at least one ε_0, $\inf\Delta = 0$ and thus it holds for all $\varepsilon < \varepsilon_0$. □

4.18 Remark. It is known from calculus that the space of all real-valued continuous functions defined on $\mathbb{R}^n$ is closed under the formation of main algebraic operations. What if the functions were defined on an arbitrary space (X, d)? We give here some informal discussion on this matter. Let $\mathbb{R}^X$ be the space of all real-valued functions defined on a set X and let $f, g \in \mathbb{R}^X$. Define the following.

(i) $f \pm g$ is the function such that for each point $x \in X$, $(f \pm g)(x) = f(x) \pm g(x)$.

(ii) fg is the function such that for each $x \in X$, $(fg)(x) = f(x) \cdot g(x)$.

(iii) $+\infty$ and $-\infty$ are not real numbers. Consequently, f/g is the function such that for all $x \in X$, $(f/g)(x) = f(x)/g(x)$, excluding $x \in X$ for which $g(x) = 0$. At all those values, the function f/g is either undefined or should be specified.

(iv) As a special case, any real-valued function multiplied by a real number is a real-valued function too.

(v) The associative (relative to multiplication) and distributive laws of functions relative to the addition and multiplication defined in (i) and (ii) are the corresponding consequences of these laws for real numbers.

Bearing in mind these observations, we conclude that the space $\mathbb{R}^X$ with the above operations is a commutative algebra over the field $\mathbb{R}$ with unity, and a vector lattice (which was also mentioned in

Example 7.7(ix), Chapter 1). A subset $\mathcal{C}((X,d);(\mathbb{R},\rho))$ (of $\mathbb{R}^X$) of all continuous functions is a subalgebra characterized by the following properties.

(a) $f,g \in \mathcal{C} \Rightarrow af+bg \in \mathcal{C}$, $\forall a,b \in \mathbb{R}$.
(b) $f,g \in \mathcal{C} \Rightarrow fg \in \mathcal{C}$. □

PROBLEMS

4.1 Show that if A is totally bounded then A is bounded. Give an example, where a bounded set is not totally bounded.

4.2 Prove that $\mathcal{C}$ in Remark 4.14 is indeed a subalgebra with properties (a) and (b).

4.3 Show that a continuous bounded function on a bounded interval need not be uniformly continuous.

In Problems 4.4 - 4.8 it is assumed that f and g are functions from $(\mathbb{R},d_e)$ to $(\mathbb{R},d_e)$.

4.4 Let $f:((-\infty,0),d_e) \to ((-\infty,0),d_e)$ be a function given by $f(x)=\frac{1}{x}$. Show that f is continuous. Explain why $f(x)$ is not uniformly continuous.

4.5 Let $f:A \to \mathbb{R}$ be a differentiable function such that its derivative f' is bounded over A, where A is an arbitrary (bounded or unbounded) interval. Show that f is uniformly continuous on A.

4.6 Show that if f and g are uniformly continuous on $\mathbb{R}$ and bounded then fg is uniformly continuous on $\mathbb{R}$ too. However, if f and g are uniformly continuous then fg need not be uniformly continuous.

4.7 Which of the following functions are uniformly continuous?

$a)$ $f(x) = \sin^2 x \ \ (x \in \mathbb{R})$.
$b)$ $f(x) = x^3 \cos x \ (x \in \mathbb{R})$.
$c)$ $f(x) = x \sin x \ (x \in \mathbb{R})$.
$d)$ $f(x) = \ln x \ (x \in [1, \infty)$.
$e)$ $f(x) = x^2 \ln x \ (x \in (1, 100))$.
$f)$ $f(x) = \sqrt{x}$, $x \in \mathbb{R}_+$.
$g)$ $f(x) = \sin x^2$.

4.8 Let f be a continuous function and g a uniformly continuous function on a set A such that $|f| \leq |g|$. Is f then uniformly continuous?

4.9 Show that in $(\mathbb{R}^n, d_e)$, any bounded set is also totally bounded.

4.10 Show that in $\mathbb{R}^n$, Euclidean and supremum metrics are equivalent.

4.11 Let A be a nonempty subset in a metric space (X, d). Define the function $f(x) = d(A, x) = \inf\{d(x, y) : y \in A\}$ as the distance from $x \in X$ to set A. Show that f is uniformly continuous.

4.12 Let (X, d) be a metric space and define the function $\rho : (X^2, X^2) \to \mathbb{R}_+$ as $\rho(\mathbf{x}, \mathbf{y}) = d(x_1, y_1) + d(x_2, y_2)$, where $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$. (a) Show that ρ is a metric on X^2. (b) Define the function $f(\mathbf{x}) = d(x_1, x_2)$. Show that $f : X^2 \to \mathbb{R}_+$ is uniformly continuous on (X^2, ρ).

4.13 Let (X, d) and (Y, ρ) be two metric spaces. Let function $[X, Y, f]$ be a map. Is the following argument true? "If f is not uniformly continuous, then there is an $\varepsilon > 0$ such that for each $\delta > 0$, there is a set $S_\delta \subseteq X$ with $d : S_\delta \times S_\delta \to (0, \delta)$, and

$$\rho : f_*(S_\delta) \times f_*(S_\delta) \to [\varepsilon, \infty).$$

If S_δ is finite, then we choose $\widetilde{\delta} = \min\{d(x, y) : x, y \in S_\delta\}$. It shows that δ is reducible to $\widetilde{\delta}$. Herewith we empty S_δ and thus

for all $x, y \in X, d(x, y) < \widetilde{\delta}$ and $\rho(f(x), f(y)) < \varepsilon$. Thus S_δ is infinite."

4.14 Give an example of a continuous function $[X, Y, f]$ from (X, d) to (Y, ρ) such that f is not Cauchy continuous.

4.15 Prove Proposition 4.16. *Let (X, d) and (Y, ρ) be two metric spaces and let function $[X, Y, f]$ be uniformly continuous on X. Then f is also Cauchy continuous.*

4.16 Prove the following statement. *Let (X, d) and (Y, ρ) be two metric spaces and let $[X, Y, f]$ and $[X, Y, g]$ be two continuous mappings. The subset $A = \{x \in X : f(x) = g(x)\}$ is closed.*

4.17 Recall [Example 4.13 (v)] that a function f from (X, d) to (Y, ρ) is said to *satisfy the Lipschitz condition on X* if there is a nonnegative constant C such that for each $x, y \in X$, $\rho(f(x), f(y)) \leq Cd(x, y)$. The smallest such C is called the *Lipschitz constant of f*. Show that f is uniformly continuous on X if f satisfies the Lipschitz condition.

4.18 Let (X, d) be a metric space such that d is bounded, that means $d(X) < \infty$. Is the function ρ defined on $\mathcal{P}(X) \times \mathcal{P}(X)$ as

$$\rho(A, B) = \sup\{d(x, y) : x \in A, y \in B\}$$

a metric on $\mathcal{P}(X)$? If it is not, show which properties of a metric hold and which do not.

NEW TERMS:

continuity at a point, 117
continuous function on a set, 117
continuity criteria in metric spaces 119, 120
Heine, Heinrich Eduard, 120
Heine's continuity criteria in metric spaces, 120
topology generated by a metric, 120
Dirichlet function, 121
equivalent metrics, 121
diameter of a set, 122
bounded (d-bounded) set, 122
d-bounded set, 122
unbounded set, 122
totally bounded set, 122
uniformly continuous function, 123
isometry between two metric spaces, 124
Lipschitz condition, 125
Lipschitz constant, 125
Cauchy continuity of a function, 125
algebra of functions, 127
$\mathcal{C}((X, d); (\mathbb{R}, \rho))$, continuous functions, 128
distance between two sets, 130

5. COMPLETE METRIC SPACES

In this section we are concerned with completeness of metric spaces previously introduced in Definition 3.1(iv). According to Problems 5.10 and 5.11, completeness is a stronger property of a subset than closeness. However, these two notions of a set coincide if the metric space is complete, as we learn it from the following theorem.

5.1 Theorem. *Let (X, d) be a complete metric space. Then, a subspace (A, d) is complete if and only if A is closed.*

Proof. Let A be closed and let $\{x_n\} \subseteq A$ be any Cauchy sequence. Because (X, d) is complete, there is a point $x \in X$ such that $\lim_{n\to\infty} x_n = x$. Then, by Corollary 3.4, $x \in A$. Thus, (A, d) is complete. Now, let (A, d) be complete and $\{x_n\}$ be any convergent sequence in A. Then this sequence is also a Cauchy sequence and hence A contains its limit. Therefore, A is closed, again, by Corollary 3.4. □

The reader should be aware of the differences between the notions of completeness and closeness of a subspace. (See Problem 5.4.)

5.2 Theorem. *A metric space (X, d) is complete if and only if every nested sequence $\{C(x_n, r_n)\}$ of closed balls, with $r_n \downarrow 0$ as $n \to \infty$, has a nonempty intersection.*

Proof. Because $r_n \downarrow 0$, for any $\varepsilon > 0$, there is an integer ν such that $r_\nu < \frac{1}{2}\varepsilon$. Given that $k > n \geq \nu$,

$$C(x_k, r_k) \subset C(x_n, r_n) \subseteq C(x_\nu, r_\nu)$$

and, consequently,

$$d(x_k, x_n) \leq 2r_\nu < \varepsilon.$$

Therefore, $\{x_n\}$ is a Cauchy sequence.

Now, assume that (X, d) is complete. Then $\{x_n\}$ converges to a point, say $x \in X$. In as much as each ball $C(x_n, r_n)$ contains the tail

$$\{x_n, x_{n+1}, \dots\}$$

of the sequence $\{x_n\}$ and because it is closed, it must contain x. Thus, $\bigcap_{n=1}^{\infty} C(x_n, r_n)$ contains x and hence it is not empty.

Now, let any nested sequence of closed balls have a nonempty intersection and let $\{x_k\}$ be a Cauchy sequence in X. By Definition 3.1 (iii), it implies the existence of an increasing subsequence $\{\nu_1, \nu_2, \dots\}$ of indices of $\{x_k\}$ such that for each n,

$$d(x_s, x_{\nu_n}) < \tfrac{1}{2^{n+1}}, \text{ for } s > \nu_n.$$

We show that the sequence $\left\{C_n = C\left(x_{\nu_n}, \tfrac{1}{2^n}\right)\right\}$ is nested. Indeed, let $y \in C_{n+1}$. Then

$$d(y, x_{\nu_{n+1}}) \leq \tfrac{1}{2^{n+1}} \quad \text{and} \quad d(x_{\nu_n}, x_{\nu_{n+1}}) < \tfrac{1}{2^{n+1}}.$$

Therefore,

$$d(y, x_{\nu_n}) < \tfrac{1}{2^n},$$

which yields that y is an interior point of C_n and thus $C_n \supset C_{n+1}$. Because by our assumption, the intersection $\bigcap_{n=1}^{\infty} C_n \neq \emptyset$, there is at least one point, say x that belongs to all balls. Furthermore, because the sequence $\{r_n\}$ of their radii is convergent to zero, the subsequence $\{x_{\nu_n}\}$ of their centers must converge to $x \in X$ and thus, by Problem 3.9, $\{x_k\}$ also converges to x. □

5.3 Remark. In the final phrase of the last theorem, x is a unique point of the intersection $\bigcap_{n=1}^{\infty} C_n$. Theorem 5.4 is a useful refinement of this statement due to Georg Cantor. Because of its similarity to Theorem 5.2, its proof is suggested as an exercise (Problem 5.4). □

5.4 Theorem (Cantor). *Let (X, d) be a complete metric space and let $\{A_n\}\downarrow \subseteq X$ be a sequence of nonempty closed subsets with*

$$\lim_{n\to\infty} d(A_n) = 0.$$

Then $\bigcap_{n=1}^{\infty} A_n$ *consists of exactly one element.* □

5.5 Definition. A function $[X, (Y, \rho), f]$ is called *ρ-bounded* if Y is a vector space and there is a nonnegative real number M such that $\rho(f(x), \mathbf{0}(x)) \leq M, \ \forall x \in X,$ where $\mathbf{0}$ is the function identically equal to $\theta \in Y$ (the origin of Y). □

5.6 Examples.

(i) Let X be a nonempty set, (Y, ρ) a metric vector space, and let $\mathcal{F}_* = \mathcal{F}_*(X; (Y, \rho))$ be the set of all ρ-bounded functions from X to Y. For all $f, g \in \mathcal{F}_*$ define

$$d_\infty(f, g) = \sup\{\rho(f(x), g(x)) : x \in X\}.$$

It can be shown (Problem 5.5) that d_∞ is a metric on $\mathcal{F}_*,$ called a *uniform* (or *supremum*) *metric.* Consequently, the convergence in $(\mathcal{F}_*, d_\infty)$ is called the *uniform convergence.* A subset of functions $\mathcal{F} \subseteq \mathcal{F}_*$ is said to be *uniformly bounded on* X if $\mathcal{F}$ is d_∞-bounded; that is, $\operatorname{diam} \mathcal{F} = \sup\{d_\infty(f, g) : f, g \in \mathcal{F}\} \leq M$ (a positive real number).

We show that any Cauchy sequence in $(\mathcal{F}_*, \rho)$ is uniformly bounded. We will make use of Problem 5.5. Let $\{f_n\}$ be a Cauchy sequence in $(\mathcal{F}_*, d_\infty)$. Therefore, for $\varepsilon = 1$, there is an $N = N(1)$ such that $d_\infty(f_n\,, f_k) < 1, \ n, k \geq N$. Let $k = N(1)$. Then,

$$d_\infty(f_n, \mathbf{0}) \leq d_\infty(f_n\,, f_N) + d_\infty(f_N, \mathbf{0}) < 1 + M(f_N),$$

where $M(f_N)$ is a "ρ-bound" of function f_N. If $M(f_i)$ is a bound of $f_i,$ then $M,$ defined as

$$M = \max\{M(f_1), \ldots, M(f_{N-1}), 1 + M(f_N)\},$$

d_∞-dominates the whole sequence $\{f_n\}$. By Problem 5.5, we have that $\{f_n\}$ is d_∞-bounded.

(ii) Assume that (Y, ρ) is a complete metric vector space. Let us show then that $(\mathcal{F}_*, d_\infty)$ is also complete. Consider a Cauchy sequence $\{f_n\} \subseteq (\mathcal{F}_*, d_\infty)$. It is obvious that for each fixed $x \in X,$

the sequence $\{f_n(x)\}$ is also Cauchy in (Y, ρ). Because (Y, ρ) is by our assumption complete, the "pointwise limit" of $\{f_n\}$ exists. Denote it by f. In other words,

$$\lim_{n \to \infty} \rho(f_n(x), f(x)) = 0, \ \forall x \in X.$$

We need to show that $f \in (\mathcal{F}_*, d_\infty)$. Because $\{f_n\}$ is a Cauchy sequence, according to (i) it is uniformly bounded by a real number M. Thus we have

$$\begin{aligned} \rho(f(x), \mathbf{0}(x)) &\leq \rho(f(x), f_n(x)) + \rho(f_n(x), \mathbf{0}(x)) \\ &\leq \rho(f(x), f_n(x)) + d_\infty(f_n, \mathbf{0}) \leq M, \end{aligned}$$

that is,

$$\rho(f(x), 0(x)) \leq \rho(f(x), f_n(x)) + M.$$

The last inequality holds for every $x \in X$ if $n \to \infty$, which yields

$$\rho(f(x), \mathbf{0}(x)) \leq M, \text{ for all } x \in X.$$

Consequently, $d_\infty(f, \mathbf{0}) \leq M$ and hence, $f \in (\mathcal{F}_*, d_\infty)$.

We only showed that

$$f_n(x) \xrightarrow{\rho} f(x),$$

for each $x \in X$, and that $f \in \mathcal{F}_*$. The assertion $f_n \xrightarrow{d_\infty} f$ is subject to Problem 5.7. □

PROBLEMS

5.1 Using similar arguments as in Example 5.6, show that the limit of any uniformly convergent sequence of continuous bounded functions from (X, d_0) to (Y, ρ) is a bounded and continuous function.

5.2 Prove Cantor's Theorem 5.4.

5.3 Let $\{C_n\}$ be a sequence of closed balls in $(\mathbb{R}^n, d_e)$ such that each of the balls C_n is centered at a point $x_0 \in \mathbb{R}^n$ and has radius $\frac{1}{n}$, $n = 1, 2, \ldots$. Find the intersection $\bigcap_{n=1}^{\infty} C_n$.

5.4 Show that if a metric space (X, d) is not complete then a closed subspace (A, d) need not be complete either. (*Hint*: Consider the metric space in Problems 2.9 and 2.10.)

5.5 Show that d_∞, defined in Example 5.6(i), is a metric on $\mathcal{F}_*$.

5.6 Let $\mathcal{F} \subseteq \mathcal{F}_*(X; (Y, \rho))$, where Y is a vector space. Prove that $\mathcal{F}$ is d_∞-bounded if and only if there is a positive constant M such that for all $f \in \mathcal{F}$, $d_\infty(f, \mathbf{0}) \leq M$.

5.7 Show that in Example 5.6 (ii) $f_n \overset{d_\infty}{\to} f$.

5.8 We can make use of the fact that the Euclidean and uniform (supremum) metrics are equivalent (see Problem 4.10) to show completeness of $(\mathbb{R}^n, d_e)$. For $n = 1$, it is well known from calculus. Prove completeness of $(\mathbb{R}^n, d_e)$ for an arbitrary n. (See Problem 3.5.)

5.9 Let (X, d) be a metric space. A subset $A \subseteq X$ is said to be of the *first category* if it can be represented as a countable union of nowhere dense sets. Otherwise, A is of the second category. Prove Baire's category theorem: *a complete metric space is of the second category*.

5.10 Let (X, d) be a metric space and $A \subseteq X$ be a complete subset. Show that A is closed.

5.11 Show that a closed set in a metric space need not be complete.

NEW TERMS:

completeness criteria, 132
Cantor's Theorem on intersection of closed sets, 133
ρ-bounded function, 134
bounded function, 134
uniform (supremum) metric, 134
supremum (uniform) metric, 134
uniform convergence continuous bounded functions, 135
uniformly bounded set of functions, 136
d_∞-bound of a function, 136
Baire's category theorem, 136

6. COMPACTNESS

Compactness is one of the key concepts in real analysis. We develop it in the present section for metric spaces and then revisit it in Chapter 3 for general topological spaces. It stems from the fact known in $\mathbb{R}$ that every bounded sequence has a convergent subsequence, which implies that any sequence in a closed bounded interval has a subsequence convergent to a point in this interval. In a general metric space, a subset A, in which every sequence has a subsequence convergent to a point in A, is called *sequentially compact* or just *compact*. Although compactness and sequential compactness are generally distinct notions in topological spaces (and they are defined in a different way), they are equivalent in metric spaces as Theorem 6.3 states.

Continuous functions defined on compact sets are uniformly continuous; continuous images of compact sets are compact (hence, closed and bounded) and this means that in normed vector spaces (introduced in the next section), continuous functions on compact sets reach their maximum values. Further applications lead to the celebrated Ascoli and Ascoli - Arzela theorems [Dsh2].

6.1 Definitions.

(i) A family of sets $\{A_i : i \in I\} \subseteq (X, d)$ is called a *cover* of a set $A \subseteq X$ if

$$A \subseteq \bigcup_{i \in I} A_i.$$

Any subfamily of $\{A_i : i \in I\}$, that covers A is called a *subcover* of A. If $\{A_i : i \in I\}$ is a family of open sets, then the corresponding cover (or subcover) is called an *open cover* (or an *open subcover*).

(ii) A set $A \subseteq (X, d)$ is called *compact* if any open cover of A has within itself a finite subcover of A, or we also say that "any open cover of A can be reduced to a finite subcover of A." Correspondingly, (X, d) is a *compact metric space* if so is X. Notice that any finite subset is compact. Consequently, to avoid triviality, in all theorems below we assume that underlying sets of spaces are infinite.

(iii) A set $A \subseteq (X, d)$ is called a *Lindelöf set* if any open cover of A contains a countable subcover of A (or can be reduced to a countable subcover). (X, d) is called a *Lindelöf space* if X is a Lindelöf set. A Lindelöf space was named after the Finnish topologist Ernst Leonard Lindelöf (1870-1946). □

A noteworthy property of Euclidean spaces is given in the following classical result.

6.2 Theorem (Lindelöf). *Any subset of $\mathbb{R}^n$ is Lindelöf set in $(\mathbb{R}^n, d_e)$. In particular, $(\mathbb{R}^n, d_e)$ is a Lindelöf space.* □

(See Problem 6.7.)

6.3 Theorem. *For a subset $A \subseteq (X, d)$, the following statements are equivalent.*

(i) *A is compact.*

(ii) *Every infinite subset of A has an accumulation point in A* (in this case A is called *Bolzano-Weierstrass compact*).

(iii) *Every sequence in A has a subsequence that converges in A* (A is called *sequentially compact*). □

Sequential compactness of a subspace implies its completeness. (See Problem 6.6.)

The proofs of the above statements are left for the reader (Problem 6.8).

Definition 6.4. A metric space is called *separable* if it has a dense countable subset. □

Example 6.5. The Euclidean metric space $(\mathbb{R}, d_e)$ is separable. A relevant dense countable subset of $\mathbb{R}$ would be $\mathbb{Q}$, the set of rational numbers. Another example is the n-dimensional Euclidean metric space with the countable, dense subset $\mathbb{Q}^n$. □

Theorem 6.6. *Any compact metric space is separable.*

Proof. Let X be compact. It is easy to see that for each $n \in \mathbb{N}$, X can be covered by the family of open balls centered at every $x \in X$ with radius $\frac{1}{n}$. Because (X, d) is compact, this open cover can be re-

duced to a finite subcover, such that $\bigcup_{x\in F_n} B(x, \frac{1}{n})$ contains X, where $F_n = \{x_1^n, \ldots, x_{k_n}^n\}$. Denote $F = \bigcup_{n=1}^{\infty} F_n$, which obviously is a countable subset of X. We show that F is dense in X, that is, $\overline{F} = X$. It is sufficient to prove that, for each $y \in X$ and $r > 0$, the open ball $B(y, r)$ contains at least one point of the set F, that is, y is a closure point of F. Taking some y and r we choose any n such that $\frac{1}{n} < r$. Then if

$$y \in X \subseteq \bigcup_{x\in F_n} B(x, \tfrac{1}{n}),$$

there is a point $x_{j_n}^n \in F_n$ such that $y \in B(x_{j_n}^n, r)$. This implies that $d(x_{j_n}^n, y) < r$ and, therefore, $x_{j_n}^n \in B(y, r)$. Consequently, $B(y, r) \cap F_n \neq \emptyset$ and $B(y, r) \cap F \neq \emptyset$. The proof of the statement is complete. □

The following two theorems belong to central results in analysis.

Theorem 6.7. *Let $A \subseteq (X, d)$ be compact. Then A is closed and bounded.*

Proof.

(i) We show that A is bounded. Obviously A is covered by the family of open balls $\{B(x, 1)\colon x \in A\}$. Since A is compact, this cover can be reduced to a finite subcover, that is, $A \subseteq \bigcup_{k=1}^{h} B(x_k, 1)$ for some integer h. Let $M = \max\{d(x_i, x_j) : i, j = 1, \ldots, h\}$. Then M is finite. For any $x, y \in A$, there are x_i and x_j such that $x \in B(x_i, 1)$ and $y \in B(x_j, 1)$. The following holds due to the triangle inequality.

$$d(x, y) \leq d(x, x_i) + d(x_i, x_j) + d(x_j, y) < 1 + M + 1 < \infty.$$

Therefore, A is bounded.

(ii) We show that A is closed, that is, that $A = \overline{A}$. Let $x \in \overline{A}$. By Theorem 3.3, there exists a sequence $\{x_n\} \subseteq A$ such that $x_n \to x$. By Theorem 6.3, if A is compact, every sequence $\{x_n\} \subseteq A$ has a subsequence that converges in A. By Problem 3.7, such a subse-

quence must have the same limit as $\{x_n\}$, that is, $x \in A$. Therefore, A is closed. □

The history of what today is called the *Heine–Borel theorem* (one of the most cited theorems in analysis) begins in the second half of nineteenth century in the work of the German mathematician Peter Gustav Dirichlet. The key issue was the concept of uniform continuity and the theorem stating that every continuous function on a closed interval is uniformly continuous. In his proof, Dirichlet implicitly used the existence of a finite subcover of a given open cover of a closed interval. He presented this proof in his lectures taught in 1862, which were published only in 1904. The first known proof of the theorem was rendered in 1872 by Émile Borel and not by Heinrich Eduard Heine (who, by the way, was a student of Dirichlet). However, Borel's formulation was restricted to countable covers and he used techniques similar to those that Heine used to prove that continuous functions on closed intervals are uniformly continuous. Consequently, Heine's name was attached due to the similarity in Heine's and Borel's approaches. Henri Lebesgue in 1898, among a few others, generalized the Heine - Borel theorem to arbitrary covers.

Theorem 6.8 (Heine - Borel). *A set $A \subseteq (\mathbb{R}^n, d_e)$ is compact if and only if A is closed and bounded.*

Proof.

(i) If A is compact, it is closed and bounded as a special case of Theorem 6.7.

(ii) If $A \subseteq (\mathbb{R}^n, d_e)$ is closed and bounded, $d_e(x, y) \leq M < \infty$, $\forall\ x, y \in A$. Let y and $\mathbf{a} = (a_1, \ldots, a_n)$ be two elements of A. Then we have:

$$|a_i| = \sqrt{a_i^2} \leq \sqrt{\sum_{k=1}^{n} (a_k - 0)^2} = d_e(\mathbf{a}, \theta)$$

$$\leq d_e(\mathbf{a}, y) + d_e(y, \theta) \leq M + d_e(y, \theta),\ i = 1, \ldots, n,$$

where θ denotes the origin of $\mathbb{R}^n$. Because each y has a finite distance to the origin, every other point of A, like $\mathbf{a}$, has also a finite

distance to the origin bounded by $M + d_e(y, \theta)$. Note that even though $d_e(\mathbf{a}, \theta) < \infty$, $d_e(\mathbf{a}, \theta)$ in general, need not have a uniform bound, unless A were bounded, which is assumed. Now we show that any d_e-bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

The steps below represent an appropriate selection procedure. Let $\{x_k\} \subseteq A$. Then $\{x_k^i\} \subseteq \mathbb{R}$ is a bounded sequence of i-coordinates (the ith-component sequence), $i = 1, \dots, n$. A bounded sequence need not to converge but does have a convergent subsequence.

For $i = 1$, let such a subsequence be $\{x_{r_1}^1, x_{r_2}^1, \dots\}$ with the limit point $x^{(1)}$. Select from the second-component sequence, the subsequence with the same indices $\{x_{r_1}^2, x_{r_2}^2, \dots\}$. This subsequence is also bounded and hence contains a convergent subsequence, say $\{x_{k_1}^2, x_{k_2}^2, \dots\}$, with a limit point $x^{(2)}$, so that the set of indices $\{k_1, k_2, \dots\} \subseteq \{r_1, r_2, \dots\}$. If we return to the subsequence $\{x_{r_1}^1, x_{r_2}^1, \dots\}$ and select single out the subsequence $\{x_{k_1}^1, x_{k_2}^1, \dots\}$, then this subsequence is also convergent and has the same limit $x^{(1)}$.

We can continue this process by taking the third-component sequence, selecting the subsequence $\{x_{k_1}^3, x_{k_2}^3, \dots\}$ and from this sequence a convergent subsequence with a limit point $x^{(3)}$. Then the above first- and second-component subsequences will be reduced to the those with the indices from the third selection and so on. Let $x = (x^{(1)}, \dots, x^{(n)})$ be the limit of the selected subsequence of $\{x_k\}$. In as much as A is closed, x must belong to A.

Therefore, we proved that an arbitrary sequence in A has a convergent subsequence in A, implying that A is sequentially compact. By Theorem 6.3, A is then compact. □

6.9 Remark. The second part of the Heine-Borel theorem does not hold for general metric spaces. That is, if A is closed and bounded, it need not be compact. For example, let X be an infinite set and let d be the discrete metric (which is finite) on X. Then X is closed and bounded.

Now consider

$$X \subseteq \bigcup_{x \in X} B(x, 1).$$

Because each of the balls covers just one point, the open cover $\{B(x,1)\colon\ x \in X\}$ cannot be reduced to a finite subcover. Therefore, X is not compact. □

The following theorem is very important in many forthcoming applications.

6.10 Theorem. *Let $f : (X,d) \to (Y,\rho)$ be a continuous function and let K be a compact subset of X. Then the image $f_*(K)$ is compact. In short, the image of a compact set under a continuous function is compact.*

Proof. Take any open cover $\{O'_i : i \in I\}$ of $f_*(K)$ to have $f_*(K) \subseteq \bigcup_{i\in I} O'_i$. Then,

$$K \subseteq f^*(f_*(K)) \subseteq \bigcup_{i\in I} f^*(O'_i).$$

Because f is continuous, $f^*(O'_i)$ is open, and because K is compact, there is a finite subcover of K by sets $f^*(O'_i)$, without loss of generality indexed by $1,\dots,n,$ that is,

$$K \subseteq \bigcup_{k=1}^{n} f^*(O'_k).$$

Because by Problem 3.6(a, b), Chapter 1, maps preserve inclusions and unions, we have

$$f_*(K) \subseteq \bigcup_{k=1}^{n} f_*(f^*(O'_k)) \subseteq \bigcup_{k=1}^{n} O'_k.$$

The last inclusion is due to Problem 2.6, Chapter 1. □

6.11 Remark. Let $f : (X,d) \to (\mathbb{R}, d_e)$ be a continuous map and let $A \subseteq (X,d)$ be compact. Then, by Theorem 6.10, $f(A)$ is compact in $(\mathbb{R}, d_e)$. By Theorem 6.7, $f(A)$ is then closed and bounded, which means that the diameter of $f(A)$ equals some $M < \infty$. As mentioned in part (ii) in the proof of the Heine-Borel Theorem, this implies that all points of $f(A)$ have a finite distance (i.e., are bounded by some M_0) to the origin, or equivalently, $|f(x)| \le M_0$, for all

$x \in A$. We have therefore shown that a continuous real-valued map on a compact set assumes a minimum and a maximum value. □

6.12 Examples.

(i) In $(\mathbb{R}, d_e)$, $\mathbb{R}$ is closed but not bounded. Therefore, by the Heine-Borel theorem, $\mathbb{R}$ is not compact.

(ii) Take as $A \subseteq (\mathbb{R}, d_e)$ the set $(0, 1]$ which is bounded but not closed and therefore is not compact. We use a different argument to prove noncompactness of $(0, 1]$. Consider the open cover of A given by the family of sets $\{(\frac{1}{n},\ 2) : n = 1, 2, \dots\}$. Obviously,

$$\bigcup_{n=1}^{\infty} (\tfrac{1}{n}, 2) = (0, 2) \supseteq A.$$

It is not possible to select any finite subcover of A, for no finite subcover would include the point 0. Yet another argument that A is not compact (due to Theorem 6.3(ii)) is that the sequence $\{\frac{1}{n}\}$ does not converge in A. □

A continuous function need not be uniformly continuous, unless it is defined on a compact set, which is a widely referred to result known for Euclidean spaces.

6.13 Theorem. *Let f: $(X, d) \to (Y, \rho)$ be a continuous function and let (X, d) be compact. Then f is uniformly continuous on X.*

Proof. Let f be continuous at x. Then, for each $\varepsilon > 0$, there is a $\delta_x > 0$, such that

$$\rho(f(x), f(y)) < \tfrac{\varepsilon}{2}$$

for all y with $d(x, y) < \delta_x$. Because X is compact, after reduction, there is an n-tuple of open balls such that

$$X \subseteq \bigcup_{i=1}^{n} B(x_i, \delta_{x_i}/2).$$

Let $\delta = \frac{1}{2} \min\{\delta_{x_1}, \dots, \delta_{x_n}\}$ and let x, y be such that $d(x, y) < \delta$. Then $x \in B(x_i, \delta_{x_i}/2)$ implies that

$$d(x, x_i) < \tfrac{1}{2}\, \delta_{x_i}$$

and

$$d(x_i, y) \le d(x, y) + d(x, x_i) < \delta + \tfrac{1}{2}\,\delta_{x_i} \le \delta_{x_i}.$$

Thus, y belongs to the ball $B(x_i, \delta_{x_i})$. Because y and x_i are within the distance of δ_{x_i}, due to continuity of f at x_i, given ε,

$$\rho(f(x_i), f(y)) < \tfrac{\varepsilon}{2}.$$

Obviously, $d(x, x_i) < \delta_{x_i}$ yields $\rho(f(x_i), f(x)) < \frac{\varepsilon}{2}$ and, therefore,

$$\rho(f(x), f(y)) \le \rho(f(x_i), f(x)) + \rho(f(x_i), f(y)) < \varepsilon. \qquad \square$$

Here is a very useful criterion of compactness.

6.14 Theorem. *A metric space (X, d) is compact if and only if it is complete and totally bounded.*

Proof.

(i) Let (X, d) be compact. Then by Problem 6.6, it is complete. Because $X \subseteq \bigcup\limits_{x \in X} B(x, \varepsilon)$ for some $\varepsilon > 0$, by compactness, the cover $\{B(x, \varepsilon)\colon x \in X\}$ can be reduced to a finite subcover, which implies total boundedness.

(ii) Let (X, d) be complete and totally bounded. We show that (X, d) is sequentially compact, which, by Theorem 6.3, would imply compactness. Let $\{x_n\}$ be a sequence in X. We construct a Cauchy subsequence.

X is totally bounded. Therefore, it can be covered by finitely many open balls of radius 1. Then, at least one of the balls, for instance B_1, contains infinitely many terms, say $\{x_k^1\}$, of this sequence. Furthermore, we cover X by balls of radius $\frac{1}{2}$ and again an infinite subsequence $\{x_k^2\} \subseteq \{x_k^1\}$ (because B_1 will also be covered) is contained in one of the balls, which we label B_2, and so on.

The desired Cauchy sequence is formed by the selection of the first term from each subsequence. Indeed, by the construction, x_1^1 and x_1^2 belong to ball B_1. Thus, $d(x_1^1, x_1^2) < 1$. Then, x_1^3 and x_1^2 belong to

ball B_2, which implies that $d(x_1^2, x_1^3) < \frac{1}{2}$, and so on. Because (X, d) is assumed to be complete, this Cauchy sequence must be convergent, thereby yielding sequential compactness of (X, d). □

PROBLEMS

6.1 Show that if $\{x_k\} \subseteq (\mathbb{R}^n, d_e)$ with $d(x_k, 0) \leq 3$, then $\{x_k\}$ has a convergent subsequence.

6.2 Define for each pair $A, B \in (X, d)$, $d(A, B) = \inf\{d(a, b) : a \in A, b \in B\}$. Let A be compact. Show that $\forall B \subseteq X$, $\exists$ an $x \in A$ such that $d(x, B) = d(A, B)$.

6.3 Let $A, B \subseteq (X, d)$ such that A is compact and B is closed. If $A \cap B = \emptyset$, show that $d(A, B) > 0$.

6.4 Let $A \subseteq (X, d)$. Show that if A is totally bounded then $\overline{A}$ is also totally bounded.

6.5 Generalize Theorem 6.6: *Any Lindelöf metric space is separable.*

6.6 Show that sequential compactness of a subspace implies its completeness.

6.7 Prove Theorem 6.2. *Any subset of $\mathbb{R}^n$ is Lindelöf set in $(\mathbb{R}^n, d_e)$. In particular, $(\mathbb{R}^n, d_e)$ is a Lindelöf space.*

6.8 Prove Theorem 6.3: *For a subset $A \subseteq (X, d)$, the following statements are equivalent.*

- (i) *A is compact.*
- (ii) *Every infinite subset of A has an accumulation point in A* (in this case A is called *Bolzano-Weierstrass compact*).
- (iii) *Every sequence in A has a subsequence that converges in A* (A is called *sequentially compact*).

6.9 Show that a mapping $f : (X, d) \to (Y, \rho)$ is continuous if and only if, for any compact subset K of X, $\mathrm{Res}_K f$ is continuous.

6.10 Show that a closed subset of a compact metric space is compact.

6.11 Let $\mathcal{K}$ be the family of all compact subsets of a metric space (X, d). List those set operations which are contained within $\mathcal{K}$. Give counterexamples for all set operations, which are not valid in $\mathcal{K}$.

NEW TERMS:

cover, 138
subcover, 138
open cover, 138
open subcover, 138
compact set, 138
compact metric space, 138
Lindelöf set, 139
Lindelöf space, 139
Lindelöf, Ernst Leonard, 139
Lindelöf theorem, 139
compactness criteria (metric space), 139
Bolzano - Weierstrass compactness (metric space), 139
sequential compactness (metric space), 139
separable metric space, 139
separability criterion (metric space), 139
compactness criterion (metric space), 140
Dirichlet, Peter Gustav, 141
Borel, Émile, 141
Heine, Heinrich Eduard, 141
Lebesgue, Henri, 141
Heine - Borel Theorem, 141
compact set under a continuous function (metric space), 143
uniform continuity criterion in compact space, 144
compactness criterion (metric space), 145
separability of a Lindelöf metric space, 146

7. NORMED VECTOR SPACES

We have already mentioned that the Euclidean metric defines the length of a vector in n-dimensional Euclidean vector space. The following generalizes the notion of vector length in a vector space and reconciles it with the notion of a special metric defined on a vector space (initially discussed in Section 5).

It is hard to say when exactly normed vector spaces were born. In 1912, the Austrian mathematician Eduard Helly published an important paper about bounded functionals on the Banach space $C_{[a,b]}$ followed by his habilitation work (submitted in 1918) and his seminal paper of 1921, "Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten", *Monatshefte fr□Mathematik und Physik* , **31**, 60 - 91. Here he considers subspaces of $\mathbb{C}^{\mathbb{N}}$ and defines a *norm* on one such subspace X, although he neither calls it the "norm" nor does he use the common notation $\|\cdot\|$. However, he uses instead the word *Abstandfunktion* (distance function) $D : X \to \mathbb{R}_+$ such that, with $x \in X$, also $\alpha x \in X$ $(\alpha \in \mathbb{F})$, with $D(\alpha x) = |\alpha| D(x)$, with $x, y \in X$, also $x + y \in X$ and $D(x + y) \leq D(x) + D(y)$, and finally, $D(x) = 0$ implies that $x = \theta$. Although norms in various forms had been introduced earlier, Helly was the first one to utilize the relationship between the norm and convexity. Shortly thereafter, in 1922, Stefan Banach defended his thesis (submitted in 1920) in which a well structured concept of the norm in abstract vector spaces has been laid out. And not only that: Banach also considered complete normed vector spaces, which Maurice Fréchet in 1928 suggested be named *Banach spaces*.

7.1 Definition. Let (X, d) be a metric space such that X is a vector space over a field $\mathbb{F}$ (which is $\mathbb{R}$ or $\mathbb{C}$). The metric d is said to be:

a) translation invariant if for all $a,\ x,\ y \in X,\ d(x + a, y + a) = d(x, y)$.

b) homothetic if for all $\alpha \in \mathbb{F}$ and $x, y \in X,\ d(\alpha x, \alpha y) = |\alpha| d(x, y)$.

If a metric d is translation invariant and homothetic we abbreviate it by TIH. □

If d is a metric on a vector space X, then we are able to measure the length of vectors, thus comparing them by setting the distance from any point $x \in X$ to one fixed point of X, the origin. If, in addition, d is TIH then we can use the properties of X as a vector space, and in some particular cases, employ even geometry, thereby emulating the Euclidean space and preserving the generality needed in applications.

7.2 Definition. Let d be a TIH metric on a vector space X, with the origin θ, over $\mathbb{F}$ (assuming that $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$). Then for all $x \in X$, we call the distance $d(x, \theta)$ the *norm of vector* x and denote it by $\|x\|$. We also call $\|\cdot\|$ the *norm on* X *induced by the TIH metric* d. The pair $(X, \|\cdot\|)$ will be referred to as a *normed vector space* (*NVS*). □

7.3 Theorem. *Let* $\|\cdot\|$*be a norm on* X *in Definition 7.2. Then, the following properties of* $\|\cdot\|$ *hold true.*

(i) $\|x\| = 0 \quad \Leftrightarrow \quad x = \theta.$

(ii) $\|\alpha x\| = |\alpha|\,\|x\|, \forall \alpha \in \mathbb{F},\ \forall\, x \in X.$

(iii) $\|x + y\| \leq \|x\| + \|y\|,\ \forall\, x, y \in X.$

(iv) *The norm* $\|\cdot\|$ *is a continuous mapping from* $(X, \|\cdot\|)$ *to* $(\mathbb{R}_+, |\cdot|)$.

Proof. Property (i) is obvious.

(ii) $$\begin{aligned}\|\alpha x\| &= d(\alpha x, \theta) = d(\alpha x, \alpha\theta) \\ &= |\alpha| d(x, \theta) = |\alpha| \|x\|.\end{aligned}$$

(iii) $$\begin{aligned}\|x + y\| &= d(x + y, \theta) = d(x, -y) \\ &\leq d(x, \theta) + d(\theta, -y) = \|x\|\end{aligned}$$

$$+ |-1|\|y\| = \|x\| + \|y\|.$$

(iv) From inequality (1.2c) we have

$$|d(x,\theta) - d(z,\theta)| = |\|x\| - \|z\|| \leq d(x,z) = \|x - z\|. \quad (7.3)$$

Thus, if $f(x) := \|x\|$, we have from Definition 4.1 that $f : (X, \|\cdot\|) \to (\mathbb{R}, |\cdot|)$ is continuous at x if for each $\varepsilon > 0$, there is a $\delta > 0$ such that $|f(x) - f(z)| < \varepsilon$ for all $z \in X$ with the property that $\|x - z\| < \delta$. Now, letting $\varepsilon = \delta$ we have the latter hold due to inequality (7.3). □

7.4 Remark. Conversely, if $\|\cdot\|$ is a real-valued nonnegative function defined on a vector space X and has properties $(i\text{-}iii)$ of Theorem 7.3, then $\|\cdot\|$ generates a TIH metric on X by setting $d(x, y) = \|x - y\|$ (show it, see Problem 7.6). □

7.5 Definition. If d in Definition 7.2 is a TIH pseudometric, then the function $\|\cdot\|$ is called a *seminorm* and correspondingly, the pair $(X, \|\cdot\|)$ is called a *seminormed vector space* (*SNVS*). □

It is easy to show that the Euclidean metric d_e on $\mathbb{R}^n$ is TIH. The associated norm induced by d_e is called the *Euclidean norm* and it is denoted $\|\cdot\|_e$.

A very important class of NVS's is introduced below.

7.6 Definition. An NVS is called a *Banach space* if it is complete with respect to the metric induced by the norm (or the norm induced by a TIH metric, as per Remark 7.4). □

7.7 Examples.

(i) The NVS $(\mathbb{R}^n, \|\cdot\|_e)$ over the field $\mathbb{R}$ with $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$ is a Banach space with the Euclidean norm (see Problem 7.1).

(ii) The NVS l^p $(p \geq 1)$ over the field $\mathbb{C}$ with the norm $\|x\|_p = [\sum_{n=1}^{\infty} |x_n|^p]^{1/p}$ is a Banach space. Observe that $\|\cdot\|_p$ indeed defines a norm (called the l^p *norm*). (See Problem 7.5.) Now let $\{x^{(n)}\}$ be a Cauchy sequence. Then as a Cauchy sequence it is

bounded (see Problem 3.4), say, by some $M \in \mathbb{R}_+$. Let $x = (x_1, x_2, \dots)$ be the pointwise limit of the sequence $\{x^{(n)}\}$. This limit exists, because each x_i is the limit of the ith-component sequence in $(\mathbb{C}, d_e)$ which is complete. First we need to show that x is an element of l^p (that is, $\|x\|_p < \infty$) and secondly, that

$$x^{(n)} \xrightarrow{l^p} x$$

(i.e., $\{x^{(n)}\}$ converges to x in l^p norm).

We have

$$\left[\sum_{k=1}^{r} |x_k|^p\right]^{1/p} = \left[\sum_{k=1}^{r} \left|x_k - x_k^{(n)} + x_k^{(n)}\right|^p\right]^{1/p}$$

(by Minkowski's inequality, Problem 1.7 (a),

$$\left[\sum_{i=1}^{n}(a_i + b_i)^p\right]^{1/p} \leq \left[\sum_{i=1}^{n} a_i^p\right]^{1/p} + \left[\sum_{i=1}^{n} b_i^p\right]^{1/p}$$

with $a_k = x_k - x_k^{(n)}$ and $b_k = x_k^{(n)}$)

$$\leq \left[\sum_{k=1}^{r} \left|x_k - x_k^{(n)}\right|^p\right]^{1/p} + \left[\sum_{k=1}^{r} \left|x_k^{(n)}\right|^p\right]^{1/p}$$

$$\leq \left[\sum_{k=1}^{r} \mid x_k - x_k^{(n)} \mid^p\right]^{1/p} + \left\|x^{(n)}\right\|_p$$

$$\leq \left[\sum_{k=1}^{r} \left|x_k - x_k^{(n)}\right|^p\right]^{1/p} + M.$$

Now, letting $n \to \infty$, we have

$$\left[\sum_{k=1}^{r} |x_k|^p\right]^{1/p} \leq M,$$

which holds for all $r = 1, 2, \dots$. Hence, with $\|x\|_p \leq M$, we showed that $x \in l^p$.

The proof that $x^{(n)} \to x$ in l^p norm is left for an exercise. (See Problem 7.7.) So, l^p is complete and therefore it is a Banach space.

(iii) Let $\mathcal{F}_*(X)$ be the space of all bounded functions on X valued in $(\mathbb{R}, d_e)$ or $(\mathbb{C}, d_e)$. Obviously, $\mathcal{F}_*$ is a vector space. The norm $\|f\|_\infty = \sup\{|f(x)| : x \in X\}$ is called the *supremum norm.* The space $(\mathcal{F}_*(X), \|\cdot\|_\infty)$ is then an NVS. The convergence in such a space (i.e., in supremum norm) is referred to as the *uniform convergence*. Now, $(\mathcal{F}_*(X), \|\cdot\|_\infty)$ is a Banach space (see Problem 7.4).

(iv) Consider $\mathcal{C}^n_{[a,b]}$ as the space of all n-times differentiable real-valued functions on a compact interval $[a, b]$. It is easily seen that $\mathcal{C}^n_{[a,b]}$ is a vector space. We introduce the following norm in $\mathcal{C}^n_{[a,b]}$.

$$\|f\|_\Sigma = \|f\|_\infty + \|f'\|_\infty + \cdots + \left\|f^{(n)}\right\|_\infty$$

Clearly, $\|\cdot\|_\Sigma$ is a norm in $\mathcal{C}^n_{[a,b]}$. We show that $\mathcal{C}^n_{[a,b]}$ is a Banach space under this norm. Let $\{f_k\}$ be a $\|\cdot\|_\Sigma$-Cauchy sequence. Then, for every $\varepsilon > 0$, there is a positive integer N such that $\forall k, j \geq N$,

$$\| f_j - f_k \|_\infty + \| f_j' - f_k' \|_\infty + \cdots + \| f_j^{(n)} - f_k^{(n)} \|_\infty < \varepsilon,$$

which implies

$$\left\| f_j^{(i)} - f_k^{(i)} \right\|_\infty < \varepsilon, \ i = 0, 1, \dots, n.$$

Therefore, by the well-known theorem from calculus (cf. Theorem 4.2, p. 508, in [Fish]), there exists a function $g_i : [a, b] \to \mathbb{R}$ to which the sequence $\{f_j^{(i)} : j = 1, 2, \dots\}$ converges uniformly and g_i is continuous, $i = 0, 1, \dots, n$. On the other hand, it holds that

$$f_k^{(i-1)}(x) - f_k^{(i-1)}(a) = \int\limits_{[a,x]} f_k^{(i)}(u)du,$$

$$i = 1, \ldots, n,\ k = 1, 2, \ldots\ .$$

Let $k \to \infty$ in the above equation in the sense of the pointwise convergence. Because $f_k^{(i)}$ and g_i are bounded functions, by the Lebesgue dominated convergence theorem (see forthcoming Chapter 6 for a rigorous proof of this theorem), we may interchange the limit and the integral and have

$$g_{i-1}(x) - g_{i-1}(a) = \int\limits_{[a,x]} g_i(u)du,\ i = 1, \ldots, n.$$

Consequently, we conclude that g_{i-1} is differentiable on $[a, b]$ and $g'_{i-1}(x) = g_i(x)$. Thus $g_0 \in \mathcal{C}^n_{[a,b]}$ implying that $\|f_k - g_0\|_\Sigma \to 0$ and $\mathcal{C}^n_{[a,b]}$ is a Banach space.

(v) The set $\mathbb{Q}$ of all rational numbers with the Euclidean norm $\|\cdot\|_e$ is an NVS, but it obviously fails to be Banach.

(vi) In Example 1.3(iv) we introduced the vector space $\mathcal{C}_{[a,b]}$ of all continuous functions on interval $[a, b] \subseteq \mathbb{R}$ and defined the supremum metric

$$d(x, y) = \sup\{|x(t) - y(t)|\colon t \in [a, b]\},$$

which induces the supremum norm

$$\|x\|_\infty = \sup\{|x(t)|\colon t \in [a, b]\}.$$

Now, the space $\left(\mathcal{C}_{[a,b]}, \|\cdot\|_\infty\right)$ is Banach. Indeed, let $\{x_n\}$ be a Cauchy sequence. Then, for each $\varepsilon > 0$, there is an $N(\varepsilon)$-tail of $\{x_n\}$. Thus, for each $t \in [a, b]$, $|x_m(t) - x_n(t)| < \varepsilon$, meaning that for each $t \in [a, b]$, $\{x_n(t)\}$ is Cauchy in $(\mathbb{R}, \|\cdot\|_e)$. Because the latter is complete, we conclude that $x_n(t)$ converges pointwise to some function $x(t)$. Consequently, $|x_m(t) - x_n(t)| < \varepsilon$ for all $m, n \geq N(\varepsilon)$ (with $m \to \infty$ for each $t \in [a, b]$) implies that

$$|x(t) - x_n(t)| \leq \varepsilon,\ \text{holding}\ \forall\, t \in [a, b]$$

and thus

$$\|x_n - x\|_\infty \leq \varepsilon.$$

The latter means that $\{x_n\}$ converges to x uniformly on $[a,b]$ and therefore, from calculus we know that x is continuous. Hence, $\left(\mathcal{C}_{[a,b]}, \|\cdot\|_\infty\right)$ is Banach. [The result also follows from (iv)].

(vii) Let $R^1_{[a,b]}$ denote the space of all real-valued Riemann integrable functions on interval $[a,b]$. Obviously, $R^1_{[a,b]}$ is a real vector space. Introduce for an element $x \in R^1_{[a,b]}$, the function

$$\|x\| = \int_a^b |x(t)|dt, \tag{7.7}$$

which obeys all properties of a norm except that $\|x\| = 0$ does not imply that $x(t) = 0$ for all $t \in [a,b]$. This is because there are integrable functions that can differ from x on infinite sets, even uncountable sets. (For the reader familiar with analysis, a function y can differ from zero function $\mathbf{0}$ on a continuum set of Lebesgue measure zero and it must be almost everywhere continuous. We discuss these and other aspects of integration in Euclidean spaces in Chapter 6, Section 3.) Thus the function defined in (7.7) is a seminorm on $R^1_{[a,b]}$.

As we have done for metrics, we can convert $\|\cdot\|$ to a norm and $\left(R^1_{[a,b]}, \|\cdot\|\right)$ to an NVS by turning to the quotient NVS space $\left(R^1_{[a,b]}|\mathrm{E}_{\|\cdot\|}, \|\!|\cdot|\!\|\right)$, where $\|\!|\cdot|\!\|$ (notice the bold face of $\|\cdot\|$) is the restrictor of $\|\cdot\|$. Obviously, an element of $R^1_{[a,b]}|\mathrm{E}_{\|\cdot\|}$ will be an equivalence class $[x]$ of all integrable functions that are equal "almost everywhere on $[a,b]$." □

The following property of NVS's asserts that the summation of their elements and multiplication by a scalar are two continuous functions with respect to the norm. This property makes NLS's an important subclass of "topological vector spaces."

7.8 Proposition. *Let X be an NVS. Then, the functions*

(i) $\varphi : X \times X \to X$ *defined as* $\varphi(x,y) = x + y$

(ii) $\psi : \mathbb{F} \times X \to X$ *defined as* $\psi(\alpha, x) = \alpha x$

are continuous.

Proof. (i) By Theorem 4.4, φ is continuous if and only if for every convergent sequence $(x_n, y_n) \to (x, y)$, the sequence $\varphi(x_n, y_n) \to \varphi(x, y)$, and this obviously is the case, because

$$\|(x_n + y_n) - (x + y)\|$$

$$\leq \|x_n - x\| + \|y_n - y\| \to 0 \text{ as } n \to \infty.$$

The proof of (ii) is similar. □

We close this section with the theorem playing a major role in the next topic on finite-dimensional NVS's.

7.9 Theorem. *Given a linearly independent set* $\{x_1, \ldots, x_n\} \subseteq (X, \|\cdot\|)$, *there is a positive constant* C *such that, for any* n*-tuple of scalars* $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$)*, the following inequality holds.*

$$\|\textstyle\sum_{i=1}^n \alpha_i x_i\| \geq C \sum_{i=1}^n |\alpha_i|. \tag{7.9}$$

Proof. Note that (7.7) holds trivially if $\sum_{i=1}^n |\alpha_i| = 0$. We thus assume that at least one $\alpha_i \neq 0$. Then, with $\beta_i := \alpha_i / \sum_{i=1}^n |\alpha_i|$, we have

$$\|\textstyle\sum_{i=1}^n \beta_i x_i\| \geq C, \text{ such that } \sum_{i=1}^n |\beta_i| = 1, \tag{7.9a}$$

which is equivalent to (7.9) if at least one of β_i's is not zero. That is, we need to show that given a linearly independent set of vectors $x_1, \ldots, x_n$, there is a positive constant C such that for any an n-tuple of scalars $\beta_i, \ldots, \beta_n$ with $\sum_{i=1}^n |\beta_i| = 1$ the inequality

$$\|\textstyle\sum_{i=1}^n \beta_i x_i\| \geq C$$

holds. Assume the opposite, that is, for any n-tuple of linearly independent vectors $x_1, \ldots, x_n$, any n-tuple of scalars $\beta_i, \ldots, \beta_n$ with $\sum_{i=1}^n |\beta_i| = 1$, and any positive constant C,

$$\|\textstyle\sum_{i=1}^n \beta_i x_i\| < C.$$

Therefore, for any $k = 1, 2, \ldots,$ there are n-tuples of linearly independent vectors $x_1, \ldots, x_n$ and an n-tuple of scalars $\beta_1^{(k)}, \ldots, \beta_n^{(k)}$ with $\sum_{i=1}^n \left|\beta_i^{(k)}\right| = 1$ such that

$$\left\|\sum_{i=1}^n \beta_i^{(k)} x_i\right\| < \tfrac{1}{k}.$$

Then,

$$\left\|\sum_{i=1}^n \beta_i^{(k)} x_i\right\| \to 0 \text{ as } k \to \infty.$$

Now, the sequence $\left\{\beta_1^{(k)}, \ldots, \beta_n^{(k)} : k \in \mathbb{N}\right\}$ is obviously bounded. Thus, by Bolzano-Weierstrass theorem, the sequence $\left\{\beta_1^{(k)} : k \in \mathbb{N}\right\}$ has a convergent subsequence $\left\{\beta_1^{(k_{s^1})} : s^1 \in \mathbb{N}\right\}$ whose limit we denote by δ_1. Then, from the subsequence $\left\{\beta_2^{(k_{s^1})} : s^1 \in \mathbb{N}\right\}$ there is a convergent subsequence $\left\{\beta_2^{(k_{s^2})} : s^2 \in \mathbb{N}\right\}$ whose limit is δ_2. Continuing sifting every next subsequence we finally arrive at the subsequences

$$\left\{\beta_i^{(k_{s^n})}, i = 1, \ldots, n : s^n \in \mathbb{N}\right\} \to \{\delta_1, \ldots, \delta_n\},$$

such that

$$\left\|\sum_{i=1}^n \beta_i^{(k_{s^n})} x_i\right\| \to 0 \text{ as } s^n \to \infty$$

and

$$\sum_{i=1}^n \left|\beta_i^{(k_{s^n})}\right| = 1$$

and thus by continuity of the Euclidean norm,

$$\sum_{i=1}^n |\delta_i| = 1. \tag{7.9b}$$

On the other hand, by continuity of norm $\|\cdot\|$ (Theorem 7.3(iv)),

$$\left\| \sum_{i=1}^{n} \beta_i^{(k_{s^n})} x_i \right\| \to \| \sum_{i=1}^{n} \delta_i x_i \| = 0.$$

Thus $\sum_{i=1}^{n} \delta_i x_i = \theta$. Because by (7.9b) not all δ_i's are zeros, $x_1, \ldots, x_n$ must be linearly dependent, which is a contradiction. □

PROBLEMS

7.1 Show that $(\mathbb{R}^n, \ \| \cdot \|_e)$ defined in Example 7.7 (i) is an NVS and then show that it is a Banach space.

7.2 Define the space l^∞ as the set of all bounded sequences $x = \{x_1, x_2, \ldots\} \subseteq \mathbb{C}$. Show that l^∞ is an NVS with the norm defined as $\|x\| = \sup\{|x_i|\colon\ i = 1, 2, \ldots\}$.

7.3 Define the space $\mathrm{c} \subseteq l^\infty$ as the subset of all convergent subsequences and let $\mathrm{c}_0 \subseteq \mathrm{c}$ be the set of all sequences convergent to zero. Show that c and c_0 are normed vector subspaces of l^∞ with the same norm as that in Problem 7.2.

7.4 Let $\mathcal{F}_*(\Omega)$ be the space of all bounded real-valued functions on Ω. Show that $\mathcal{F}_*$ is a vector space. Let $\|f\|_\infty = \sup\{|f(\omega)|\colon\ \omega \in \Omega\}$ be the supremum norm defined in Example 7.7(iii). Show that the supremum norm in $\mathcal{F}_*$ is indeed a norm and show that $\mathcal{F}_*$ is a Banach space with respect to this norm.

7.5 Show that $\| \cdot \|_p$ in Example 7.7(ii) is a norm.

7.6 Let $\| \cdot \|$ be a real-valued nonnegative function defined on a vector space X over a field $\mathbb{F}$ (which is $\mathbb{R}$ or $\mathbb{C}$) and let it have properties $(i\text{-}iii)$ of Theorem 7.3. Show that $\| \cdot \|$ generates a TIH metric on X by $d(x, y) = \|x - y\|$.

7.7 Show that the pointwise limit x of the sequence $\{x^{(n)}\}$ in Example 7.7(ii) is also an l^p-limit.

NEW TERMS:

Helly, Eduard, 149
Banach, Stefan, 149
Fréchet, Maurice, 149
translation - invariant metric, 149
homothetic metric, 149
TIH metric, 150
norm, 150
normed vector space (NVS), 150
NVS, 151
seminorm, 151
seminormed vector space (SNVS), 151
SNVS, 151
Euclidean norm, 151
Banach space, 151
l^p-norm, 151
$(\mathcal{F}_*(X), \|\cdot\|_\infty)$, space of all bounded functions, 153
supremum norm, 153
uniform convergence, 153
Σ-norm, 153
$\left(R^1_{[a,b]}, \|\cdot\|\right)$, semi-NVS of Riemann integrable functions,155
$\left(R^1_{[a,b]}|\mathrm{E}_{\|\cdot\|}, \|\cdot\|\right)$, NVS of Riemann integrable functions, 155
l^∞ space, 158

8. FINITE-DIMENSIONAL NORMED VECTOR SPACES

The notion of a finite-dimensional normed linear space was introduced by the German mathematician Hermann Minkowski (1864 - 1909) in his monograph *Geometrie der Zahlen* (Teubner, Leipzig, 1896). To honor his contribution, such spaces are often called *Minkowski spaces*.

We begin this section with the following significant property of finite-dimensional NVS's:

8.1 Theorem. *Any finite-dimensional NVS is a Banach space*.

Proof. Let n be the dimension of an NVS X and let $\{b_1, \dots, b_n\}$ be a Hamel basis for X. Let $\{x_k\} \subseteq X$ be a Cauchy sequence. Then, every element of the sequence can uniquely be expressed as

$$x_k = \sum_{i=1}^{n} \alpha_i^{(k)} b_i \tag{8.1}$$

Because $\{x_k\}$ is Cauchy, for every $\varepsilon > 0$, there is an $N(\varepsilon)$-tail $\{x_N, x_{N+1}, \dots\}$ of $\{x_k\}$ such that $\|x_k - x_m\| < \varepsilon$ for all elements of the $N(\varepsilon)$-tail. Since $b_1, \dots, b_n$ are linearly independent, due to Theorem 7.7, there is a universal constant $C > 0$ such that

$$\|x_k - x_m\| = \left\| \sum_{i=1}^{n} \left[\alpha_i^{(k)} - \alpha_i^{(m)} \right] b_i \right\| \geq C \sum_{i=1}^{n} \left| \alpha_i^{(k)} - \alpha_i^{(m)} \right|$$

good for all $k, m \geq N(\varepsilon)$.

Therefore, for all k and m greater than or equal to $N(\varepsilon)$,

$$\left| \alpha_i^{(k)} - \alpha_i^{(m)} \right| \leq \sum_{i=1}^{n} \left| \alpha_i^{(k)} - \alpha_i^{(m)} \right| < \varepsilon / C.$$

Consequently, $\left\{ \alpha_i^{(k)} \right\}$ is a Cauchy sequence in $(\mathbb{F}, |\cdot|)$ and the latter is complete (whether $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$). If α_i is the limit of $\alpha_i^{(k)}, i = 1, \dots, n,$ then the element $x := \sum_{i=1}^{n} \alpha_i b_i$ belongs to X and we have by the triangle inequality:

$$\|x - x_k\| = \left\|\sum_{i=1}^{n}\left[\alpha_i - \alpha_i^{(k)}\right]b_i\right\| \leq \sum_{i=1}^{n}\left|\alpha_i - \alpha_i^{(k)}\right|\|b_i\|.$$

For each i, $\alpha_i^{(k)}$ is convergent to α_i. Therefore, for each choice of $\frac{\varepsilon}{\|b_i\|n} > 0$, there is an $N_i\left(\frac{\varepsilon}{\|b_i\|n}\right)$-tail of $\left\{\alpha_i^{(k)}\right\}$ such that for all terms of the $N_i\left(\frac{\varepsilon}{\|b_i\|n}\right)$-tail, the inequality $\left|\alpha_i - \alpha_i^{(k)}\right| < \frac{\varepsilon}{\|b_i\|n}$ holds. If $N = \max\{N_1, \ldots, N_n\}$, then

$$\|x - x_k\| \leq \sum_{i=1}^{n}\left|\alpha_i - \alpha_i^{(k)}\right|\|b_i\| < \varepsilon \tag{8.1a}$$

for all $k \geq N$, that is, for each ε, there is an $N(\varepsilon)$-tail of $\{x_k\}$ with property (8.1a). Thus, the sequence $\{x_k\}$ is convergent. □

Combining Theorems 8.1 and 5.1 we arrive at the following.

8.2 Corollary. *Any finite-dimensional subspace of a normed vector subspace is a closed subspace of this space.* □

According to Theorem 6.7, every compact subset of a metric space is closed and bounded. The converse of this theorem was part two of Theorem 6.8 (Heine - Borel) holding for the Euclidean space. Beyond the Heine-Borel theorem, we learned that this converse did not hold in discrete metric spaces. It stands to reason to investigate in which spaces the converse of Theorem 6.7 is valid.

8.3 Definition. We say that an NVS (or metric space) has *the Heine - Borel property* if every closed and bounded subset of the space is compact. □

8.4 Theorem. *Every finite-dimensional NVS space has the Heine - Borel property.*

Proof. Suppose A is a closed and bounded subset of an n-dimensional NVS X. Let $\{x_k\} \subseteq A$ be a sequence and $\{b_1, \ldots, b_n\}$ be a Hamel basis for X. Then,

$$x_k = \sum_{i=1}^{n}\alpha_i^{(k)}b_i,\ k = 1, 2, \ldots \tag{8.3}$$

Because A is bounded, the sequence is also bounded, that is, there is a constant M such that $\|x_k\| \leq M$ for each k. Due to Theorem 7.7, there is a constant $C > 0$ such that

$$\|x_k\| = \left\|\sum_{i=1}^{n} \alpha_i^{(k)} b_i\right\| \geq C \sum_{i=1}^{n} \left|\alpha_i^{(k)}\right|. \tag{8.3a}$$

Inequality (8.3a) yields $\left|\alpha_i^{(k)}\right| \leq M$, $i = 1, \ldots, n$. By the Bolzano - Weierstrass theorem, $\left\{\alpha_i^{(k)}\right\}$ has a convergent subsequence for each k. Repeating the same process as in the proof of Theorem 7.7, we can extract from the sequence $\left\{\alpha_1^{(k)}, \ldots, \alpha_n^{(k)} : k \in \mathbb{N}\right\}$ a convergent subsequence $\left\{\alpha_1^{(k_s)}, \ldots, \alpha_n^{(k_s)}\right\}$ whose limit we denote by $\{\delta_1, \ldots, \delta_n\}$. Consequently, by (8.3), the sequence $\{x_k\}$ has a subsequence $\{x_{k_s}\}$ that can easily be shown to converge to point

$$x := \sum_{i=1}^{n} \delta_i b_i \in A,$$

in as much as A is closed. In other words, A is sequentially compact. Therefore, by Theorem 6.3, A is compact. □

The following theorem by Frigyes Riesz sheds more light on the Heine - Borel property for NVS's, where we establish a necessary and sufficient condition for it.

8.5 Theorem (F. Riesz). *An NVS $(X, \|\cdot\|)$ has the Heine - Borel property if and only if it is finite-dimensional.*

To prove this theorem we need the following.

8.6 Lemma. *Suppose $(X, \|\cdot\|)$ is an infinite-dimensional NVS. Then, there is a linearly independent sequence $\{x_n\} \subseteq S(\theta, 1)$ (where $S(\theta, 1)$ is the unit sphere centered at zero) such that for each pair x_j, x_k of distinct elements, $\|x_j - x_k\| \geq 1$.*

Proof. We use the induction principle to prove this statement. Assume that for $n \geq 1$, there is an n-tuple $\{x_1, \ldots, x_n\}$ of linearly independent elements all belonging to the unit sphere $\partial B(\theta, 1)$ such that $\|x_j - x_k\| \geq 1$. Denote $A = \text{span}\{x_1, \ldots, x_n\}$. Let $z \in A^c$ and

let $\delta = d(z, A^c)$. Because, by Corollary 8.2, A is closed, A^c is open and thus there is a ball $B(z, r) \subseteq A^c$ implying that $\delta > 0$.

Now, since A is a subspace, it follows that $\frac{1}{\delta}A = \left\{\frac{1}{\delta}x : x \in A\right\} = A$. Denote $z_0 = z/\delta$. Because the metric induced by the norm is homothetic, we have

$$1 = \tfrac{1}{\delta}d(z, A) = d(z_0, A).$$

From the latter it follows that there is a sequence $\{y_n\} \subseteq A$ such that

$$\|y_n - z_0\| \to 1.$$

(Indeed, $1 = d(z_0, A) = \inf\{\|z_0 - x\| : x \in A\}$ means that for each $\varepsilon > 0$, there is an $x \in A$ such that $1 \leq \|z_0 - x\| < 1 + \varepsilon$.)

Now, we have

$$\|y_n\| \leq \|y_n - z_0\| + \|z_0\| < R + \|z_0\| = : M.$$

Therefore, $\{y_n\} \subseteq C(\theta, M) \cap A$, which is a closed and bounded subset of A. Because A has the Heine-Borel property, $C(\theta, M) \cap A$ is compact and thus by Theorem 6.3, $\{y_n\}$ has a convergent subsequence $\{y_{n_k}\}$ that has a limit point, say y, in A. Because $\|y_n - z_0\| \to 1$, we must also have $\|y_{n_k} - z_0\| \to 1$ and thus

$$1 = d(z_0, A) \leq \|y - z_0\|$$

$$\leq \|y - y_{n_k}\| + \|y_{n_k} - z_0\| \to 0 + 1$$

implying that $\|y - z_0\| = 1$.

Define $x_{n+1} : = y - z_0$. We show that x_{n+1} fits the pattern of the elements from $\{x_1, \ldots, x_n\}$.

(i) Obviously, $x_{n+1} \in S(\theta, 1)$.

(ii) If x_{n+1} is linearly dependent with $\{x_1, \ldots, x_n\}$, then $\sum_{i=1}^n \alpha_i x_i = z_0 - y$. Because $y \in A$, clearly $y = \sum_{i=1}^n \beta_i x_i$ and hence $z_0 = \sum_{i=1}^n (\alpha_i + \beta_i)x_i$ and thus must belong to A. So, it implies the linear independence of the tuple $\{x_1, \ldots, x_{n+1}\}$.

(iii) Finally, we show that $\|x_j - x_{n+1}\| \geq 1$. In as much as A is a subspace, $A = A - y$ if $y \in A$ and thus

$$1 = d(z_0, A) = d(z_0 - y, A - y) = d(x_{n+1}, A)$$

$$\leq d(x_{n+1}, x) = \|x_{n+1} - x\|, \text{ for each } x \in A.$$

In particular, it holds true for $x_j \in \{x_1, \ldots, x_n\} \subseteq A$.

Thus, if we continue this process, we indeed collect a sequence $\{x_n\}$ as the lemma claims. □

Now we are back to the proof of the theorem.

Proof of Theorem 8.5. Suppose that X is infinitely dimensional and let $S(\theta, 1)$ be the unit sphere in X. Obviously, $S(\theta, 1)$ is bounded. Furthermore, $S(\theta, 1) = C(\theta, 1) \backslash B(\theta, 1)$, where $C(\theta, 1)$ is the unit closed ball, and thus $S(\theta, 1) = C(\theta, 1) \cap B^c(\theta, 1)$ is the intersection of two closed subsets. By Lemma 8.6, there is a sequence $\{x_n\} \subseteq S(\theta, 1)$, which due to the property $\|x_m - x_n\| \geq 1$, is not convergent, nor can it have a convergent subsequence. Thus, by Theorem 6.3, $S(\theta, 1)$ is not compact. Therefore, X fails to have the Heine - Borel property. □

8.7 Definition. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space X are called *equivalent* or *similar* (in notation, $\|\cdot\|_1 \approx \|\cdot\|_2$) if there are two positive constants $K \leq M$ such that for any $x \in X$,

$$K\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1. \qquad \square$$

We show that on a finite-dimensional space, all norms are equivalent.

8.8 Proposition. *The similarity in Definition 8.7 defines an equivalence relation on the set of all norms on X.* □

(See Problem 8.1.)

8.9 Proposition. *If $\|\cdot\|_1$ and $\|\cdot\|_2$ are two equivalent norms on a vector space X and if $\{x_n\}$ is a sequence in X, then*

$$(i) \quad x_n \overset{\|\cdot\|_1}{\longrightarrow} x \;\Leftrightarrow\; x_n \overset{\|\cdot\|_2}{\longrightarrow} x.$$

(ii) $\{x_n\}$ is Cauchy in $\|\cdot\|_1$ if and only if $\{x_n\}$ is Cauchy in $\|\cdot\|_2$.

Proof. Because $\|\cdot\|_1 \approx \|\cdot\|_2$, there is a positive constant M such that

$$\|x\|_2 \leq M\|x\|_1. \tag{8.9}$$

(i) Let $x_n \overset{\|\cdot\|_1}{\to} x$. Then, for any $\varepsilon > 0$, there is an $N(\varepsilon)$ tail $\{x_N, x_{N+1}, \dots\}$ of $\{x_n\}$ such that $\|x - x_n\| < \varepsilon/M$ for all $x_n \in \{x_N, x_{N+1}, \dots\}$. From (8.9) we have

$$\|x - x_n\|_2 \leq M\|x - x_n\|_1 < M\varepsilon/M = \varepsilon.$$

Obviously, the roles of the norms can be interchanged.

(ii) If $\{x_n\}$ is $\|\cdot\|_1$-Cauchy, then for each $\varepsilon > 0$ there is an $N(\varepsilon)$ tail $\{x_N, x_{N+1}, \dots\}$ of $\{x_n\}$ such that $\|x_m - x_n\| < \varepsilon/M$ for all $x_m, x_n \in \{x_N, x_{N+1}, \dots\}$. From (8.9) we have

$$\|x_m - x_n\|_2 \leq M\|x_m - x_n\|_1 < M\varepsilon/M = \varepsilon.$$

Obviously, the roles of the norms can be interchanged. □

8.10 Corollary. *If $\|\cdot\|_1$ and $\|\cdot\|_2$ are two equivalent norms on a vector space X, then $(X, \|\cdot\|_1)$ is Banach if and only if $(X, \|\cdot\|_2)$ is Banach.* □

8.11 Example. In n-dimensional Euclidean space the norms

$$\|x\|_e = \sqrt{x_1^2 + \cdots + x_n^2}$$

and

$$\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\},$$

where $x = (x_1, \dots, x_n)'$, are equivalent. Indeed, from

$$n\|x\|_\infty^2 \geq x_1^2 + \cdots + x_n^2$$

we have

$$\|\cdot\|_\infty \geq \tfrac{1}{\sqrt{n}}\|\cdot\|_e.$$

On the other hand, obviously, $\|\cdot\|_\infty \leq \|\cdot\|_e$. Thus, with the choice of $K = \frac{1}{\sqrt{n}}$ and $M = 1$ we have

$$K\|\cdot\|_e \leq \|\cdot\|_\infty \leq M\|\cdot\|_e. \qquad \square$$

Example 8.11 can be generalized in

8.12 Theorem. *In a finite-dimensional vector space all norms are equivalent*.

Proof. Let $\{b_1, \ldots, b_n\}$ be a Hamel basis for X and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be any two norms on X. Then, if $x \in X$ and thus is representable as $x = \sum_{i=1}^n \alpha_i b_i$, by Theorem 7.9, there is a positive constant C_1 such that

$$\|x\|_1 = \|\textstyle\sum_{i=1}^n \alpha_i b_i\|_1 \geq C_1 \sum_{i=1}^n |\alpha_i|.$$

Let $m := \max\{\|b_i\|_2 : i = 1, \ldots, n\}$. Then,

$$\|x\|_2 \leq \textstyle\sum_{i=1}^n |\alpha_i| \|b_i\|_2 \leq m \sum_{i=1}^n |\alpha_i| \leq M\|x\|_1,$$

where $M = \frac{m}{C_1}$. Analogously,

$$\|x\|_1 \leq \tfrac{1}{K}\|x\|_2. \qquad \square$$

8.13 Definition. Let X be an n-dimensional vector space over $\mathbb{F}$ (which is $\mathbb{R}$ or $\mathbb{C}$) with a Hamel basis $\{b_1, \ldots, b_n\}$. For any $x \in X$ such that $x = \sum_{i=1}^n \alpha_i b_i,\ \alpha_i \in \mathbb{F},$ define

$$\|x\|_* = \sqrt{\sum_{i=1}^{n} |\alpha_i|^2}. \tag{8.13}$$

According to Problem 8.2, this is a norm on $(X, \mathbb{C})$, called the *standard norm*. □

Notice that the standard norm is also a norm on $\mathbb{F}^n$.

8.14 Remarks.

(i) Proposition 8.9(i) reconciles the notions of equivalence between two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ and two metrics d_1 and d_2 induced by $\|\cdot\|_1$ and $\|\cdot\|_2$ (see Remark 7.4) in any NVS. Indeed, according to Remark 4.9, any two equivalent norms generate the same topology in an NVS, i.e. $\tau(\|\cdot\|_1) = \tau(\|\cdot\|_2)$. If in addition, an NVS is finite-dimensional, then by Theorem 8.12, any two norms induce the same topology. Furthermore, any norm in a finite-dimensional NVS induces the same topology as that of the standard norm.

(ii) The standard norm in (8.13) clearly resembles the n-dimensional Euclidean norm, which according to (i) induces a unique topology on $(\mathbb{R}, \|\cdot\|)$, with respect to any norm $\|\cdot\|$, called the *usual topology*, in notation $(\mathbb{R}, \tau_e)$. □

PROBLEMS

8.1 Prove Proposition 8.8. *The similarity in Definition* 8.7 *defines an equivalence relation on the set of all norms on X.*

8.2 Prove that the function $\|\cdot\|_*$ in Definition 8.13 is a norm.

8.3 Define the function $f : \mathbb{F}^n \to \mathbb{R}_+$ as

$$f(\alpha_1, \ldots, \alpha_n) = \|\textstyle\sum_{i=1}^{n} \alpha_i b_i\|.$$

Prove that f is continuous with respect to $\|\cdot\|_e$.

Problems 8.4 - 8.6 offer an alternative proof to Theorem 8.12 by showing that any norm $\|\cdot\|$ on a finite-dimensional NVS X (which is assumed throughout) is equivalent to the standard norm $\|\cdot\|_*$.

8.4 Let $(X, \|\cdot\|)$ be a finite-dimensional NVS and $\|\cdot\|_*$ be the standard norm defined in (8.13). Show that there is a positive constant M such that for each $x \in X$,

$$\|x\| \leq M\|x\|_*.$$

8.5 Let $(X, \|\cdot\|)$ be a finite-dimensional NVS. Let $S(0,1)$ denote the unit sphere in $\mathbb{F}^n$ centered at zero and f be the function on $\mathbb{F}^n$ defined in Problem 8.3. Show that f reaches its maximum and minimum values in $\mathbb{R}_+$ on $S(0,1)$.

8.6 Let $(X, \|\cdot\|)$ be a finite-dimensional NLS. Show that there is a positive K constant such that

$$K\|\cdot\|_* \leq \|\cdot\|.$$

8.7 Let $\|\cdot\|_1, \ldots, \|\cdot\|_k$ be norms in $\mathbb{R}^n$ and let $\|\cdot\| = \sum_{i=1}^k \|\cdot\|_i$. Show that $\|\cdot\|$ is a norm in $\mathbb{R}^n$ and that $\|\cdot\|$ induces a topology, identical to the usual topology.

8.8 Let Y be a finite-dimensional subspace of an NVS $(X, \|\cdot\|)$. Suppose there is a bounded sequence $\{y_n\} \subseteq Y$. Show that this sequence has a subsequence convergent in Y. (This generalizes a known result from calculus that any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.)

NEW TERMS:

Minkowski, Herman, 160
Minkowski spaces, 160
Heine -Borel property, 161
Riesz, Frigyes, 162
Riesz criterion of Heine - Borel property, 162
equivalent (similar) norms, 164
similar (equivalent) norms, 164
$\|x\|_*$, standard norm, 167
$(\mathbb{R}, \tau_e)$, usual topology, 167

Chapter 3

Elements of Point-Set Topology

1. TOPOLOGICAL SPACES

In Definition 4.5, Chapter 2, we called the collection of all open sets $\tau(d)$ of a metric space (X, d) the *topology induced by a metric.* We recall that this collection of open sets or topology is closed with respect to the formation of arbitrary unions and finite intersections. We understand that the topology of a metric space carries the main information about its structural fingerprint. For instance, equivalent metrics possess the same topology. In addition, through the topology we could establish the continuity of a function (see Theorem 4.6, Chapter 2) without need of a metric. This all leads to an idea of defining a structure more general than distance on a set, a structure that preserves convergence and continuity.

Mathematics historians are not in complete agreement about the roots of topology and who should get full credit for being its initiator. Most consider that topology, as the theory of structures, has its basis in the work of the German mathematician Felix Hausdorff, who published his fundamental monograph, *Grundzüge der Mengenlehre* (*Principles of Set Theory*), in Leipzig in 1914. It was "immediately" preceded by Maurice Fréchet's 1906 pioneering introduction to metric spaces. (Notice that contemporary topology has branched out into several specialized areas, such as *general topology*, *algebraic topology*, and *combinatorial topology*. The very topology founded by Hausdorff was what we now refer to as *general topology*, also called *point-set topology*, which is deeply bound to classical analysis.) Nicolas Bourbaki [Bou2], regarded German Bernhard Georg

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0_3, © Springer Science+Business Media, LLC 2013

Riemann's work (his doctoral and habilitation theses and a paper on Abelian functions) from 1851 to 1857 as revolutionary and qualified him as the creator of topology, because he was the first to recognize where topological ideas were needed. In 1870, Georg Cantor (apparently inspired by Riemann's work), in connection with the representation of real-valued functions by Fourier series, was concerned with the characterization of sets on which the function's value can be altered leaving the series invariant. This yielded more advanced concepts of topological *accumulation point* (introduced earlier by Karl Weierstrass), *derived set*, *closed set*, *connected set*, *dense set* and others that further led to the topological big bang. The very word *topology* was introduced for the first time in 1836 by the German Johann B. Listing, who used this as the notion of a "new analysis."

Topology has evolved further ever since. Most of the fundamental results in general topology were developed in works by Germans Felix Hausdorff, Heinz Hopf, and Hermann Weyl, Russians Pavel Alexandrov and Pavel Urysohn, Poles Stefan Banach, Kazimierz Kuratowski, and Waclaw Sierpiński, American Eliakim H. Moore and James Alexander, and an influential Bourbaki group of mostly French mathematicians. (Nicolas Bourbaki is a pseudonym chosen by a group of young mathematicians in France in 1935 to represent the essence of a "contemporary mathematician." The founding members were all connected to the Ecole Normale Supérieure in Paris and included Henri Cartan, Claude Chevalley, Jean Coulomb, Jean Delsarte, Jean Dieudonné, Charles Ehresmann, René de Possel, Szolem Mandelbrojt and André Weil. After World War II they were joined by the Polish - American Samuel Eilenberg.)

1.1 Definition. Let $X \neq \emptyset$. A collection τ of subsets of X is called a *topology on* X or a family of *open sets*, if:

(i) $X, \emptyset \in \tau$.

(ii) $\{O_i : i \in I\} \subseteq \tau \Rightarrow \bigcup_{i \in I} O_i \in \tau$.

(iii) τ is $\cap$-*stable*, that is, $O_1, O_2 \in \tau \Rightarrow O_1 \cap O_2 \in \tau$.

[Observe that property (iii) implies inductively that the intersection of any finite collection of open subsets will also be open.] A

carrier X endowed with a topology τ is said to be a *topological space.* The topological space is denoted by (X, τ). □

1.2 Examples.

(i) Let (X, d) be a metric space and let $\tau(d)$ be the topology generated by the metric d (see Definition 4.5, Chapter 2). Due to Theorem 2.5, Chapter 2, the collection of all open sets generated by metric d contains all arbitrary unions and finite intersections. Moreover, $\emptyset$ and X are also open, so that $\tau(d)$ is indeed a topology as defined above. We call it a *metric topology*. For instance, the topology in $\mathbb{R}^n$ generated by the Euclidean metric d_e was called the *usual* (or *standard* or *natural*) *topology* and it is denoted by τ_e [see Remark 8.14 (ii), Chapter 2].

(ii) Let X be a nonempty set. Then the pair $\{X, \emptyset\} = \tau_0$ is a trivial example of a topology. It is obviously the smallest topology on X, and it is called the *indiscrete topology.* Another trivial example of a topology is $\mathcal{P}(X)$, the collection of all subsets of X. This is the largest possible topology on X, and it is called the *discrete topology*.

(iii) For $A \subseteq X$, $\tau_1 = \{X, \emptyset, A\}$ is a topology "induced by set A."

(iv) Let $X = \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ be the extended real line. Let $\overline{\tau}_e \subseteq \mathcal{P}(X)$ be the following collection of sets.

$O \in \overline{\tau}_e$ if and only if

1) $O \cap \mathbb{R} \in \tau_e$.

2) $\infty \in O$ or $-\infty \in O$, then there is an $a \in \mathbb{R}$ such that $(a, \infty] \subseteq O$ or an $a \in \mathbb{R}$ such that $[-\infty, a) \subseteq O$, respectively.

Then $\overline{\tau}_e$ is a topology on $\overline{\mathbb{R}}$ (see Problem 1.1) referred to as the *extended real line topology*.

(v) Let (X, τ) be a topological space and let $Y \subseteq X$. Define the system of subsets $\tau_Y = \{O \cap Y : O \in \tau\}$. We show that τ_Y is a topology on Y. Indeed, Y and $\emptyset$ obviously belong to τ_Y. Let $\{U_i :$

$i \in I\} \subseteq \tau_Y$. Then, $\forall i \in I$, there is $O_i \in \tau$ such that $O_i \cap Y = U_i \in \tau_Y$. Now $\bigcup_{i\in I} O_i \in \tau$ and therefore, $Y \cap \bigcup_{i\in I} O_i \in \tau_Y$. On the other hand, due to the distributive law,

$$Y \cap \bigcup_{i\in I} O_i = \bigcup_{i\in I} (Y \cap O_i) = \bigcup_{i\in I} U_i \in \tau_Y.$$

It can similarly be shown that τ_Y is closed with respect to the formation of all finite intersections. Therefore, τ_Y is a topology on $Y \subseteq X$, called the *relative topology of* τ *on* Y. The pair (Y, τ_Y) is called a *subspace*. For instance, take the Euclidean metric space $(\mathbb{R}, d_e)$ and let $Y = [0, 1]$. Then the set $(\frac{1}{2}, 1]$ is open in (Y, τ_Y). □

1.3 Remarks.

(i) Let X be a nonempty set and let τ and τ' be two topologies on X. If $\tau \subseteq \tau'$, then we say τ is *weaker* (or *smaller* or *coarser*) *than* τ'. We also say that τ' is *stronger* (or *larger* or *finer*) *than* τ. As follows from Examples 1.2(ii) and (iii), $\tau_0 \subseteq \tau_1 \subseteq \mathcal{P}(X)$. The indiscrete topology is, therefore, the coarsest topology on X, whereas $\mathcal{P}(X)$ is the finest topology on X.

(ii) If (X, d) is a metric space and $\tau(d)$ is the topology induced by d (also called the *metric topology*), then $(X, \tau(d))$ is said to be a *metrizable* (*topological*) *space*. Therefore, a metrizable space is a topological space with a topology that comes from some metric. □

1.4 Definition. Let (X, τ) be a topological space. A subset $A \subseteq X$ is called τ-*closed* or just *closed* if $A^c \in \tau$. We symbolically denote τ^c the collection of all τ-closed subsets of X. Of course, τ^c is not the complement of τ with respect to $\mathcal{P}(X)$. □

As in the case of metric spaces, we can easily prove that X and Ø are closed, finite unions of closed sets are closed, and arbitrary intersections of closed sets are closed.

In Definitions 1.5 below we introduce some important notions for topological spaces. It is beneficial to support these definitions by examples immediately following the introduced notions. To reference the examples, we assign them letter D followed by the prefix of the definition.

1.5 Definitions.

(i) Let (X, τ) be a topological space. A subset $A \subseteq X$ is called a *neighborhood of a point* $x \in X$ if x belongs to some open subset of A. Specifically, if $A \in \tau$ then A is called an *open neighborhood of* x.

[*Example D1.5*(i). Let $X = \mathbb{R}$ and $\tau = \{\mathbb{R}, \emptyset, \{1\}, (3, 4], \{1\} \cup (3, 4]\}$. Then $\{1\}$ is an open neighborhood of 1. Because $[3, 5]$ contains $(3, 4]$, it is a neighborhood of $3\frac{1}{2}$, $(-2, 0)$ is not a neighborhood of -1, since there is no open neighborhood that contains -1 other than $\mathbb{R}$.]

(ii) A point x is called an *interior point* of a set A if A is a neighborhood of x. [In the above case, -1 is not an interior point of $(-2, 0)$.] The set of all points interior to A is called the *interior of* A and is denoted by $\mathring{A}$ or by *Int*(A).

[*Example D1.5*(ii). In *Example D1.5*(i), 1 is the interior point of the set $\{1\}$. The interior of set $A = [3, 5]$ is $\mathring{A} = (3, 4]$.]

(iii) The collection of all neighborhoods of a point $x \in X$ is called the *neighborhood system at* x and it is denoted by $\mathcal{U}_x$. (The letter $\mathcal{U}$ stems from the German word *Umgebung* which means *neighborhood.*) An arbitrary subcollection $\mathcal{B}_x \subseteq \mathcal{U}_x$ is called a *neighborhood base at* x (or a *fundamental system of neighborhoods of* x), if every neighborhood $U_x \in \mathcal{U}_x$ is a superset to least one $B_x \in \mathcal{B}_x$. Any element $B_x \in \mathcal{B}_x$ is called a *base neighborhood.* Clearly, $\mathcal{U}_x$ itself is a neighborhood base at x. Obviously, $\mathcal{B}_x$ is a neighborhood base at x if and only if there is another neighborhood base $\mathcal{D}_x$ such that every base neighborhood $D_x \in \mathcal{D}_x$ is a superset to at least one base neighborhood B_x from $\mathcal{B}_x$.

[*Example D1.5*(iii). Let $\{B(x, \frac{1}{n}), \quad n = 1, 2, \dots\}$ be the sequence of d_e-open balls centered at a point $x \in \mathbb{R}^n$. Clearly, it is a base neighborhood at x. Another neighborhood base at x, which contains the above neighborhood base, is the system of all open balls with rational radii, centered at x. We can also take the system of all open balls with positive real radii, centered at x. This system contains the first two neighborhood bases.] □

A neighborhood base $\mathcal{B}_x$ at x is in general a more "economical system" of neighborhoods than the whole neighborhood system $\mathcal{U}_x$; and, as it will be shown, it is as informative about the structure of the space in the vicinity of x as $\mathcal{U}_x$ is. Technically, it is of greater advantage in various proofs to use a base neighborhood than to use an arbitrary neighborhood.

As follows from the definition, an arbitrary set A need not be a neighborhood of all of its points. For instance, $[0, 1]$ is not a neighborhood for points 0 and 1 in the usual topology $(\mathbb{R}, \tau_e)$. More about the nature of neighborhoods is contained in the following propositions that the reader can easily verify.

1.6 Proposition. *$A \subseteq X$ is a neighborhood for all of its points if and only if A is open.* □

(See Problem 1.4.)

1.7 Proposition. *$\mathring{A}$ is the largest open set contained in A. Furthermore, $A = \mathring{A}$ if and only if A is open.* □

(See Problems 1.5.)

In particular, it follows that A is open if and only if $A = \mathring{A}$.

1.8 Definitions.

(i) $x \in X$ is called a *closure* (or *adherent*) *point for a set* A if any neighborhood of x has a nonempty intersection with A. We also say that any neighborhood of x *meets* A. The set of all closure points of A is called the *closure* of A and it is denoted by $\overline{A}$. (Sometimes, when working with relative topologies it is necessary to emphasize that the closure of A is with respect to the carrier X, and it is more reasonable to use the notation $Cl_X A$. However, for brevity we still use the notation $\overline{A}$, whenever X is the only carrier in mind.)

[*Example D1.8*(i). In the topology introduced in *Example D1.5*(i), let us take $A = (-2, 0)$. Then we have

$$\overline{A} = (-\infty, 1) \cup (1, 3] \cup (4, \infty), \tag{1.8}$$

and $\overline{A} = \emptyset$. Indeed, for any x belonging to sets $(-\infty, 1)$, $(1, 3]$ or $(4, \infty)$, $\mathbb{R}$ is the only neighborhood of x; thus $\mathbb{R} \cap (-2, 0) \neq \emptyset$. Observe that 1 is not a closure point of A, since $\{1\}$ is a neighborhood (of 1) such that $\{1\} \cap A = \emptyset$. For set $B = \{-1\}$ we have

$$\overline{B} = (-\infty, 1) \cup (1, 3] \cup (4, \infty) = \overline{A}. \ ]$$

(ii) A subset $A \subseteq X$ is said to be *dense in* X if $\overline{A} = X$. $A \subseteq X$ is said to be *nowhere dense* if $Int(\overline{A}) = \emptyset$.

[*Example D1.8*(ii). Consider *Example D1.8*(i). For the set $A = (-2, 0)$, from (1.8), we have

$$(\overline{A})^c = \{1\} \cup (3, 4],$$

and $Int(\overline{A}) = \emptyset$ (because, as mentioned above, $\mathbb{R}$ is the only neighborhood for all of its points), i.e., A is nowhere dense. The set

$$C = \{-1\} \cup \{1\} \cup (3, 4] \text{ is dense in } X.]$$

(iii) A point $x \in X$ is called an *accumulation point* (or *cluster point*) of a set A if every neighborhood of x contains at least one point of A other than x. The set of all accumulation points is called the *derived set* and is denoted by A'. Directly from the definitions of closure and accumulation points it follows that $A' \subseteq \overline{A}$.

[*Example D1.8*(iii). In *Example D1.8*(i), $A' = \overline{A}$.]

(iv) A point $x \in X$ is called a *boundary point* of a set A if every neighborhood of x contains at least one point of A and at least one point of A^c. Analogous to metric spaces, the set of all boundary points of A is

$$\partial A = \overline{A} \cap \overline{A^c}$$

and is called the *boundary* of A.

[*Example D1.8*(iv). In *Example D1.8*(i), $\overline{A} = \emptyset$ and $\partial A = \overline{A}$.]

(The closure of A is evidently the smallest closed set containing A; and A is closed if and only if $A = \overline{A}$. See Problem 1.6.)

(v) A topological space (X, τ) is called *separable* if there exists at most a countable dense subset of X. □

1.9 Proposition. *The closure of a set A is the smallest closed set containing A. Furthermore, A is closed if and only if $A = \overline{A}$.*

(See Problem 1.6.)

Convergence of sequences introduced below generalizes that of Section 3, Chapter 2, for metric spaces, and it is preparatory for the more general type of convergence of nets and filters treated in [Dsh2].

1.10 Definition. Let $\{x_n\colon\ n = 1, 2, \dots\} \subseteq (X, \tau)$ be a sequence and let A be a set. A subsequence $Q_N = \{x_n\colon\ n = N, N+1, \dots\}$ is called an $N(A)$-tail of $\{x_n\}$, for some $N \geq 1$, if $Q_N \subseteq A$. A sequence $\{x_n : n = 1, 2, \dots\} \subseteq X$ is said to *converge to a point* $x \in X$ if for every neighborhood U_x of x, there is an $N(U_x)$-tail of $\{x_n\}$. The point x is said to be a *limit point* of the sequence.

Notice that in the literature on topology, the notions of limit and accumulation points of a set are synonymous. We do not use both of them in the same context and invent a new term for the benefits of forthcoming constructions. A point x is said to be a *sequential limit point of a set* A if x is a limit point of some sequence $\{x_n\} \subseteq A$. □

Unlike metric spaces, a sequence in a topological space can have more than one limit as we learn from the following example.

1.11 Examples.

(i) Let $X = \mathbb{R}$, let $\tau = \{\mathbb{R}, \emptyset, (-2, 3], [-1, 2]\}$, and let $x_n = \frac{1}{n}$, $n = 1, 2, \dots$. Then, $\{\frac{1}{n}\}$ converges to all points of the set $[-1, 2]$, because for each point $x \in [-1, 2]$, its open neighborhoods are $\mathbb{R}$, $(-2, 3]$, and $[-1, 2]$, each one of which contains the whole sequence. Furthermore, for any point $x \in (-2, 3]^c$, $\mathbb{R}$ is the only neighborhood of x and $\mathbb{R}$ contains the entire sequence. Thus, $\{\frac{1}{n}\}$ converges to all points in $\mathbb{R}$. □

(ii) Let (X, τ) be a so-called *cocountable* (or *countable complement*) *topology* defined as $\tau := \{X, \emptyset, C^c\colon\ C$ is at most countable$\}$ introduced in Problem 1.9. We assume that $X = \mathbb{R}$. Consider

the sequence $\left\{1, 0, \frac{1}{2}, 0, \frac{1}{3}, 0, \dots\right\}$, which converges to 0 in the Euclidean metric or usual topology. Not so in the cocountable topology. The complement of the subsequence $\left\{\frac{1}{n}\right\}$ is obviously an open neighborhood of 0 that contains zero and does not contain a single element of $\left\{\frac{1}{n}\right\}$. Thus, $O_x = \left\{\frac{1}{n}\right\}^c$ does not contain any tail of the sequence.

Using the same argument we can show that the sequence does not converge at all. Indeed, if x were a limit of the sequence other than 0 or any other term of the sequence, then $\left\{1, 0, \frac{1}{2}, 0, \frac{1}{3}, 0, \dots\right\}^c$ would be an open neighborhood of x that does not include a single term of the sequence and thus cannot be its limit. □

Very important classes of topological spaces we make use of are introduced in the following.

1.12 Definitions.

(i) A topological space (X, τ) is said to *satisfy the first axiom of countability* (or to be *first countable*), if each point $x \in X$ has an at most countable neighborhood base.

(ii) A topological space (X, τ) is said to *satisfy the second axiom of countability* (or to be *second countable*), if (X, τ) has an at most countable base. □

1.13 Proposition. *Let (X, τ) be a first countable topological space and $x \in X$. Then, there is a sequence $\{x_n\} \subseteq X$ convergent to x.*

Proof. If $\mathcal{U}_x = \{U_1, U_2, \dots\}$ is an at most countable neighborhood base at x, we can form $B_1 = U_1$ and $B_n = \bigcap_{i=1}^{n} U_i, n = 1, 2, \dots$. Then, $\mathcal{B}_x = \{B_1, B_2, \dots\}$ is clearly a monotone nonincreasing neighborhood base at x. Each B_n is nonempty and it contains x. Using the axiom of choice, there is a sequence $\{x_n\}$ such that $x_n \in B_n$. Now, this sequence obviously converges to x by Problem 1.23. This convergence is with respect to $\mathcal{B}_x$, but it clearly is valid with respect to any other neighborhood system. □

1.14 Proposition. *Suppose (X, τ) is a first countable topological space and $A \subseteq X$. Then, $x \in \overline{A}$, if and only if there is a sequence $\{x_n\} \subseteq A$ such that $x_n \to x$. In other words, the closure of a set A contains all of its sequential limit points or equivalently, every closure point of A is a sequential limit point of A.*

Proof.

(i) Let $x \in \overline{A}$ and let $\mathcal{B}_x$ be a monotone nonincreasing countable neighborhood base (as per the construction in Proposition 1.13). By the definition of a closure point, each $B_n \in \mathcal{B}_x$ meets A. Thus, from $A \cap B_n$ we can choose an x_n thereby forming a sequence $\{x_n\} \subseteq A$. As in Proposition 1.13, $x_n \to x$.

(ii) Conversely, if there is a sequence $\{x_n\} \subseteq A$ such that $x_n \to x$, for each neighborhood U_x there is a tail of the sequence in U_x and therefore U_x meets A (through this tail). Consequently, x is a closure point. □

1.15 Remarks.

(i) Notice that in part (ii) of Proposition 1.14 we did not utilize the first countability of the space and yet were able to show that each limit point is a closure point. The converse of this (meaning that each closure point is a limit point) does require the first countability as we learn it from Example 1.16 below.

(ii) In metric spaces (see Corollary 3.4, Chapter 2), a point is a closure point of a set A if and only if it is a sequential limit point of A. This does not apply to topological spaces. More specifically, a sequential limit point is always a closure point as in Proposition 1.4 (ii), but the converse is not true. Let x be a closure point of A. If $x \in A$, then setting $x_n = x,$ we have a sequence convergent to x. If $x \notin A$ then for each neighborhood U_x of $x,$ as a closure point, $U_x \cap A \neq \emptyset$. In this case, it is not clear how to form a sequence convergent to $x,$ that is, how to ensure that for each U_x, there is an $N(U_x)$-tail, for we do not have the flexibility of metric spaces with balls like $B(x, \frac{1}{n})$ of Theorem 3.3, Chapter 2. In Example 1.16 we demonstrate a case of a topology where a set A contains all of its sequential limit points and yet is not closed, or, in other words, some closure points of A are not its sequential limit points. □

1.16 Examples.

(i) Let (X, τ) be a *cocountable topology* introduced earlier in Example 1.11(ii) as τ: $= \{X, \emptyset, C^c : C$ is at most countable$\}$ on an infinite set X. We assume that X is uncountable. Let $\{x_n\}$ be a sequence with infinitely many distinct terms, that is, as a set $\{x_n\} \approx \mathbb{N}$. Thus, there exists a subsequence $\{x_{n_k}\}$ with all distinct terms. Suppose there is a limit x of $\{x_n\}$. Take $O_x = \{x_{n_k}\}^c \cup \{x\}$. Then, O_x is an open neighborhood of x that does not contain any tail of $\{x_n\}$. Therefore, any such sequence cannot have a limit.

(ii) Now, under the condition of (i), let $\{x_n\}$ be a sequence with all but finitely many terms identical. Thus, for some N, $x_N = x_{N+1} = \ldots = : x$. Consequently, any open neighborhood O_x contains the N-tail of $\{x_n\}$ and thus $x_n \to x$.

It is easy to verify that any sequence in the cocountable topology can have at most one limit.

(iii) Consider the cocountable topology τ on $\mathbb{R}$ introduced in Example 1.11 (ii). Take $A = (a, b)$ where $a < b$. Let $\{x_i\} \subseteq A$ be any sequence. Then, by the definition of τ, the complement of $\{x_i\}$ is open (and disjoint from $\{x_i\}$). If this sequence has a limit $x \in A^c$, then this limit should belong to the open set $O_x = \{x_i\}^c$ (because $\{x_n\} \subseteq A \Rightarrow A^c \subseteq \{x_n\}^c$), so that O_x can serve as an open neighborhood of x.

However, this neighborhood does not have a single element of the sequence and, therefore, x cannot be its limit; or equivalently, this sequence cannot converge to any point of A^c. Therefore, if x is a limit of $\{x_n\}$, then $x \in A$, and by (ii), the sequence must have all but finitely many identical terms. Consequently, for each $x \in A$, there is a sequence in A convergent to x. So, $A = (a, b)$ contains all of its sequential limit points.

However, A is not closed. To see this, take a in association with set $A = (a, b)$. Let O_a be any open set of the form $\mathbb{R}\backslash\{$*any sequence not containing* $a\}$. Then O_a is an open neighborhood of a. Any such neighborhood O_a meets A, and thus a is an accumulation point of A (as *any neighborhood of* a *contains at least one point of* A *other than* a). Thus, A is not closed, for otherwise, by Problem 1.11, it

would contain all of its accumulation points. Now, a is thus a closure point, but not a sequential limit point of A apparently, because X fails to be first countable, according to Proposition 1.14. □

1.17 Remark. In conclusion, topological spaces fail to have the property the metric spaces do, namely,

$$x \in \overline{A} \;\Leftrightarrow\; \exists \text{ a sequence in } A \text{ whose limit is } x.$$

With no additional condition, we learned that this result was valid for metric spaces (see Theorem 3.3, Chapter 2). In other words, a set was closed if and only if it contained all of its limit points. □

PROBLEMS

1.1 Show that the collection $\overline{\tau}$ of sets introduced in Example 1.2 (iv) is a topology in $\overline{\mathbb{R}}$.

1.2 Let X be a nonempty set and $\tau = \{X, \emptyset, C^c : C \subseteq X \text{ and } C \text{ is finite}\}$. Show that τ is a topology on X. τ is called the *cofinite* (or *finite complement*) *topology* on X.

1.3 Let $X = \mathbb{R}$ and let $\tau = \{X, \emptyset, (-\infty, 1], [1, \infty), (3, 10]\}$. Is τ a topology on $\mathbb{R}$? If not, complement τ by some subsets to a topology (and be reasonable).

1.4 Prove Proposition 1.6. *$A \subseteq X$ is a neighborhood for all of its points if and only if A is open.*

1.5 Prove Proposition 1.7. (*Hint*: Show that $\overline{A}^{\square}$ contains all open sets that are contained in A and use Proposition 1.6.)

1.6 Prove Proposition 1.9. *The closure of a set A is the smallest closed set containing A. Furthermore, A is closed if and only if $A = \overline{A}$.*

1.7 Show that:

$$(a) \quad A \subseteq B \;\Rightarrow\; \overline{A} \subseteq \overline{B}.$$

(b) $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

(c) $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$; show that the inverse inclusion need not hold.

(d) $A \subseteq B \Rightarrow \mathring{A} \subseteq \mathring{B}$.

(e) $Int(A \cap B) = \mathring{A} \cap \mathring{B}$.

(f) Is it true that $\overline{Int(A)} = \overline{A}$?

1.8 Show that $\overline{A} = A \cup \partial A$.

1.9 Let X be an infinite set and let τ be a cocountable topology introduced in Example 1.11 (ii) as $\tau := \{X, \emptyset, C^c : C \text{ is at most countable}\}$. Show that τ is a topology on X.

1.10 Show that $\overline{A} = A + \partial A$ (*Hint*: Proceed in the same way as in Problem 2.11, Chapter 2, and work with a neighborhood instead of a ball.)

1.11 Prove that a subset of a topological space is closed if and only if it contains all of its accumulation points.

1.12 Show that $\overline{A} = A \cup A'$.

1.13 Let $\widetilde{\tau} = \{\mathbb{R}, (-1, 1], [0, 5), \{0\}, \{10\}\}$.

$a)$ Extend $\widetilde{\tau}$ to the smallest topology τ in $\mathbb{R}$ generated by $\widetilde{\tau}$.

$b)$ Let $A = (-7, -5]$, $B = (0, 7]$, and $C = [-\frac{1}{2}, 20)$.

Find the sets

$$\overline{A}, \overline{B}, \overline{C}, \mathring{A}, \mathring{B}, \mathring{C}, A', B', C', \partial A, \partial B, \partial C$$

and determine whether $A, B,$ and C are dense in $\mathbb{R}$.

1.14 Show that $\partial A = \emptyset$ if and only if A is open and closed.

1.15 Show that $(\overline{A})^c \subseteq \overline{A^c}$.

1.16 Show that the inverse inclusion in the previous problem holds if and only if A is closed and open.

1.17 This provides an equivalent definition of a closure point. Show that $x \in \overline{A}$ if and only if $\forall U_x \in \mathcal{U}_x,\ U_x \cap \overline{A} \neq \emptyset$.

1.18 A subset of a topological space (X, τ) is a G_δ *set* if it is a countable intersection of open sets. A *subset* F_σ is defined as a countable union of closed sets. Show that G_δ can be represented as a countable intersection of a monotone nonincreasing sequence of open sets. Likewise, show that F_σ can be represented as a countable union of a monotone nondecreasing sequence of closed sets.

1.19 Show that any neighborhood system at a point is $\cap$-stable.

1.20 Let $\mathcal{U}_x$ be a neighborhood system at x. Show that for each $U \in \mathcal{U}_x$, there is a $V \in \mathcal{U}_x$ such that $U \in \mathcal{U}_y$ for all $y \in V$.

1.21 Let $\mathcal{B}_x$ be a neighborhood base at x. Show that if $A, B \in \mathcal{B}_x$, then there is a base neighborhood $C \in \mathcal{B}_x$ such that $C \subseteq A \cap B$.

1.22 A set O is open in (X, τ) if and only if O contains base neighborhoods for all of its points.

1.23 Show the validity of the following statement. *A sequence* $\{x_n: n = 1, 2, \dots\} \subseteq X$ *converges to a point* $x \in X$ *if and only if for every base neighborhood* $B_x \in \mathcal{B}_x$ *of* x *(where* $\mathcal{B}_x$ *is an arbitrary neighborhood base), there is an* $N(U_x)$*-tail of* $\{x_n\}$.

1.24 Let $\{\tau_i : i \in I\}$ be an indexed family of topologies on a nonempty set X. Show that $\bigcap_{i \in I} \tau_i$ is a topology on X.

1.25 Let $\mathcal{S} \subseteq \mathcal{P}(X)$ be a collection of subsets of a nonempty set X. Show that there is a topology containing $\mathcal{S}$. Also show that there is the weakest topology containing $\mathcal{S}$, in notation $\tau(\mathcal{S})$.

1.26 In the context of Problem 1.18 about G_δ and F_σ sets, show that in $(\mathbb{R}^n, \tau_e)$, each $G_\delta \in \tau_e^c$ and each $F_\sigma \in \tau_e^c$.

NEW TERMS:

Hausdorff, Felix, 171
Fréchet, Maurice, 171
Bourbaki, Nicolas, 171
Riemann, Bernhard, 172
Cantor, Georg, 172
Weierstrass, Karl, 172
Listing, Johann, 172
Hopf, Heinz, 172
Weyl, Hermann, 172
Alexandrov, Pavel, 172
Urysohn, Pavel, 172
Banach, Stefan, 172
Kuratowski, Kazimierz, 172
Sierpiński, Waclaw, 172
Moore, Eliakim, 172
Alexander, James, 172
Cartan, Henri, 172
Chevalley, Claude, 172
Coulomb, Jean, 172
Delsarte, Jean, 172
Dieudonné, Jean, 172
Ehresmann, Charles, 172
Possel, René de, 172
Mandelbrojt, Szolem, 172
Weil, André, 172
Eilenberg, Samuel, 172
topology, 172
open sets, 172
intersection-stable ($\cap$ -stable) family of sets, 172
topological space 173
metric topology, 173
usual topology, 173
standard (natural) topology, 173
indiscrete topology, 173
discrete topology, 173

topology induced by a set, 173
extended real line topology, 173
relative topology (subspace), 174
subspace, 174
weaker (coarser, smaller) topology, 174
coarser topology, 174
stronger (finer, larger) topology, 174
finer topology, 174
metric topology, 174
metrizable topological space, 174
closed set, 174
τ^c, τ-closed sets, 174
neighborhood of a point, 175
open neighborhood of a point, 175
interior point, 175
$Int(A)$, interior of set A, 175
neighborhood system at a point, 175
Umgebung, 175
neighborhood base at a point, 175
fundamental system of neighborhoods at a point, 175
base neighborhood, 175
closure (adherent) point for a set, 176
neighborhood of a point that meets a set, 176
$Cl(A)$, closure of set A, 176
dense set, 177
nowhere dense set, 177
accumulation (cluster) point, 177
cluster (accumulation) point, 177
derived set, 177
boundary point of a set, 177
∂A, boundary of set A, 177
separable topological space, 178
$N(A)$-tail of a sequence, 178
convergence to a point, 178
limit point of a sequence, 178
sequential limit point of a set, 178
cocountable (countable complement) topology, 178

countable complement topology, 178
first axiom of countability, 179
first countable topological space, 179
second axiom of countability, 179
second countable topological space, 179
cofinite (finite complement) topology, 182
finite complement topology, 182
G_δ set, 184
F_σ set, 184
weakest topology containing a family of sets, 184

2. BASES AND SUBBASES FOR TOPOLOGICAL SPACES

In the previous section, we introduced the notion of a collection of open sets, called a topology. In many applications, describing an entire topology on a carrier is difficult and sometimes even impossible. This predicament is manageable if one deals instead with a smaller collection of sets, which is not a topology, but which generates a topology and thereby can be readily extended to a topology. With a similar idea, we have come to introduce neighborhood bases in place of more populous neighborhood systems for points.

Take, for example, a metric space. Although the family of all open balls does not yield a topology, every open set, as we know, can be made of the union of some collection of open balls, and consequently, it leads to a topology and gives rise to the notion of a base for a topology.

When specifying a base for a topology as its fingerprint, what is the connection between a base for the topology and neighborhood bases collected over all points? This question will also be answered in this section.

2.1 Definition. Let (X, τ) be a topological space. A subcollection $\mathcal{B}$ of open sets is called a *base for* τ if every open set is a union of some elements of $\mathcal{B}$. (Specifically, it follows that $\emptyset$ must be an element of $\mathcal{B}$.) The elements of $\mathcal{B}$ are called *base sets*. □

With no major difficulty (and with hints provided), the reader will be able to establish a very useful criterion of a base for τ, subject to Problem 2.2. An important relation between bases and neighborhood bases (see Definition $1.5(iii)$) is given in the following theorem.

2.2 Theorem. *$\mathcal{B}$ is a base for τ if and only if, $\emptyset \in \mathcal{B}$ and, for every point $x \in X$, there is a neighborhood base $\mathcal{B}_x$ consisting of open sets such that $\mathcal{B}_x \subseteq \mathcal{B}$.*

Proof. We have to show that $\mathcal{B}$ is a base for τ if and only if, for every $x \in X$ and each neighborhood U_x of x, there is a base neighborhood $B_x \in \mathcal{B}$ such that $B_x \subseteq U_x$.

(i) Let $\mathcal{B}$ be a base for τ and let U_x be a neighborhood of a point $x \in X$. Without loss of generality we assume that U_x is open. (Otherwise, take any open neighborhood $O_x \subseteq U_x$ of x and work with O_x instead.) If U_x is open, there exists a subcollection of $\mathcal{B}$ whose union equals U_x. Thus, at least one set of this subcollection, say B_x ($\in \mathcal{B}$), must contain x, and $B_x \subseteq U_x$. Observe that by Definition 1.5(iii), B_x is then an element of a neighborhood base and $\mathcal{B}_x = \{B_x\}$ forms a neighborhood base of x. Therefore, each neighborhood base $\mathcal{U}_x$ of x has at least one neighborhood base $\mathcal{B}_x$ of x such that $\mathcal{B}_x \subseteq \mathcal{B}$ and each $U_x \in \mathcal{U}_x$ is a superset of at least one $B_x \in \mathcal{B}_x$.

(ii) Let $\mathcal{B} \subseteq \tau$ and assume that for every $x \in X$, there is a neighborhood base $\mathcal{B}_x \subseteq \mathcal{B}$. Let O be an arbitrary open set. Then, by our assumption and by the definition of a neighborhood base, for any point $x \in O$ (because O is a neighborhood of x), there is a base neighborhood $B_x \in \mathcal{B}_x$ such that $B_x \subseteq O$. Thus $O = \bigcup_{x \in O} B_x$ (union of all such $B_x \in \mathcal{B}$). Hence, every open set $O \in \tau$ can be formed of a union of some elements of $\mathcal{B}$ or, equivalently, $\mathcal{B}$ is a base for τ. □

2.3 Examples.

(i) Let $\{\mathcal{B}_z : z \in X\}$ be an arbitrary collection of open neighborhood bases at all points. Then, $\bigcup_{z \in X} \mathcal{B}_z$ can be regarded as an example of a base for τ. Indeed, as in Theorem 2.2, take a point x from any open set O. Then, O is a neighborhood of x and thus it belongs to the neighborhood system at x. If $\mathcal{B}_x \in \{\mathcal{B}_z : z \in X\}$ is a neighborhood base, then there is a base neighborhood $B_x \in \mathcal{B}_x$ included in O. Collecting all such neighborhoods of all points of O, we can represent O as the union $\bigcup_{x \in O} B_x$. Hence, $\{\mathcal{B}_z : z \in X\}$ is a base for the topology τ.

(ii) As mentioned at the beginning of this section, in any metric space (X, d), the collection of all open balls is a trivial example of a base for the corresponding metrizable topological space. Indeed, by Definition 2.3, Chapter 2, for each open neighborhood O_x of $x \in X$, there exists an open ball $B(x, \varepsilon) \subseteq O_x$. Earlier [in Example D1.5(iii)], we showed that $B(x, \varepsilon)$ is a base neighborhood at x. Thus

by Theorem 2.2, the system $\{B(x,\varepsilon)\colon\ x \in X,\ \varepsilon \in (0,\infty)\}$ is a base for $\tau(d)$. As in Example D1.5(iii), a neighborhood base at x can be reduced to the system $\mathcal{B}_x = \{B(x,q)\colon\ q \in \mathbb{Q},\ q > 0\}$ of all balls with rational radii. Consequently, by Theorem 2.2, the collection of all open balls with rational radii is a base for $\tau(d)$. (Note that these balls are centered at all $x \in X,$ so consequently, this base need not be countable.)

(iii) We introduce an open rectangle in $(\mathbb{R}^n, \tau_e^n)$. More formalism on this related to the product topology is follows in Section 5. A set

$$O^{(1)} \times \cdots \times O^{(n)}$$

is called an *open rectangle* (or *parallelepiped*) in $\mathbb{R}^n$ if each $O^{(i)}$ is an open set in $\mathbb{R}$. [See also Definition 5.4(vi), Chapter 1.] An open rectangle is referred to as *simple* or *base* if each $O^{(i)}$ is an open bounded interval. Let $\mathcal{R}$ be the system of all base rectangles in $(\mathbb{R}^n, \tau_e^n)$ along with the empty set Ø and let $B(x,r)$ be an open ball centered at some point x. Then, for each r, there is a hypercube C_x centered at x (which, in particular, is a base rectangle) that can be inscribed in $B(x,r)$. Thus, the system $\mathcal{C}_x$ of all hypercubes at x can serve as a neighborhood base at x in τ_e^n and by (i), $\mathcal{C} = \bigcup\limits_{x\in\mathbb{R}^n} \mathcal{C}_x$ is a base for τ_e^n.

(iv) The collection of all singletons $\{x\} \in \mathcal{P}(X)$, along with Ø, is a base for the discrete topology on X. □

2.4 Theorem. *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be bases for topologies τ_1 and τ_2 on X. Then the following statements hold true.*

(i) *If $\mathcal{B}_1 \subseteq \tau_2$, then $\tau_1 \subseteq \tau_2$.*

(ii) *If $\mathcal{B}_1 = \mathcal{B}_2$, then $\tau_1 = \tau_2$.*

(iii) *If $\tau_1 \subseteq \tau_2$ and $\mathcal{B}_2 \subseteq \tau_1$, then $\tau_1 = \tau_2$.*

Proof.

(i) (Observe that $\mathcal{B}_1$ need not be a base for τ_2.) Indeed, by the definition of a base, each $O^1 \in \tau_1$ can be represented as $O^1 = \bigcup_i B_i$. However, because $B_i \in \tau_2$, it follows that $\bigcup_i B_i = O^1 \in \tau_2$.

(ii) By part (i), $\tau_1 \subseteq \tau_2$ and $\tau_2 \subseteq \tau_1$, and thus $\tau_1 = \tau_2$. In other words, a base uniquely defines a topology. (Note that although one topology may have different bases, a base cannot share different topologies.)

(iii) First note that it does not follow that $\mathcal{B}_2$ is a base for τ_1. In fact, $\mathcal{B}_2$ need not even be a subcollection of τ_1. However, if in addition, $\mathcal{B}_2 \subseteq \tau_1$, then by part (i), $\tau_2 \subseteq \tau_1$ and therefore, $\tau_1 = \tau_2$. Indeed, $\mathcal{B}_2 \subseteq \tau_1 \subseteq \tau_2$ implies that $\tau_1 = \tau_2$. □

When constructing a topology on a carrier, it is often beneficial to start with a collection, yet smaller and even more rudimentary than a base. With such a reduction the formation of product topologies and quick and tame continuity criteria of functions becomes even more rewarding. Recall that a function f, corresponding between two metric spaces X and Y, is continuous if and only if inverse images under f of open sets in Y are open in X. Remarkably, continuity of f can be verified frequently on the basis of much less populous families of so-called "subbase" sets in Y. This is established and elaborated in Section 4. We begin with the following.

2.5 Definition. Let $\mathcal{S} \subseteq \mathcal{P}(X)$ such that $\bigcup_{A \in \mathcal{S}} A = X$. If there exists the weakest topology on X containing $\mathcal{S}$, then it is called the *topology generated by* $\mathcal{S}$, and the collection $\mathcal{S}$ is called a *subbase on* X. (Note that $\mathcal{S}$ can directly restore only X, whereas $\mathcal{B}$ restores all open sets, including Ø. Clearly, a base $\mathcal{B}$ for a topology τ, besides τ itself, offers a trivial example of a subbase on X.) □

We justify Definition 2.5 with the following.

2.6 Proposition. *The weakest topology on X generated by a subbase exists*.

(See Problem 1.25.)

2.7 Definition. As per Proposition 2.6, if $\mathcal{S}$ is a subbase on X, then the weakest topology generated by $\mathcal{S}$ is denoted by $\tau(\mathcal{S})$ and we call $\mathcal{S}$ the *subbase for* $\tau(\mathcal{S})$. □

If we have an arbitrary collection $\mathcal{S}$ of subsets of X, it seems plausible to expand $\mathcal{S}$ by supplementing it with X, $\emptyset$, as well as all unions and finite intersections of elements of $\mathcal{S}$. Theorem 2.8 below refines this idea by showing that if $\mathcal{S}$ is a subbase, on the way to the construction of $\tau(\mathcal{S})$ we need to extend $\mathcal{S}$ with all finite intersections of its elements to arrive at a base for $\tau(\mathcal{S})$.

2.8 Theorem. *Let $\mathcal{S}$ be an arbitrary subcollection of $\mathcal{P}(X)$ such that $\bigcup_{A\in\mathcal{S}} A = X$, and let $\mathcal{B}$ be an extension of $\mathcal{S}$ that contains $\emptyset$ and all finite intersections of elements of $\mathcal{S}$. Then $\mathcal{B}$ is a base for $\tau(\mathcal{S})$.*

Proof. Let $\tau' = \left\{ \bigcup_{B\,\in\,\mathcal{B}'} B : \mathcal{B}' \subseteq \mathcal{B} \right\}$, where $\mathcal{B}$ is defined in the condition of this statement. We prove that τ' is a topology on X. It is sufficient to show that τ' contains all finite intersections; the other properties of τ' as a topology are obvious. Also, for brevity in notation, we show this for the case of the intersection of just two "open" sets. Let U and V be two elements of τ'. By the definition of $\mathcal{B}$,

$$U = \bigcup_{i\in I} U_i \ \text{ and } \ V = \bigcup_{j\in J} V_j, \ (\text{with } U_i, V_j \in \mathcal{B}),$$

where in turn

$$U_i = \bigcap_{k=1}^{n_i} S_k^i \text{ and } V_j = \bigcap_{s=1}^{m_j} S_s^j, \ (S_k^i, S_s^j \in \mathcal{S}).$$

Then

$$U \cap V = \ (\bigcup_{i\in I} \bigcap_{k=1}^{n_i} S_k^i) \cap (\bigcup_{j\in J} \bigcap_{s=1}^{m_j} S_s^j) = \bigcup_{r\in I\cup J} \bigcap_{p=1}^{r} S_p^r \ \in \tau'.$$

Now, because by the construction of τ', $\mathcal{B}$ is a base for τ', and $\mathcal{B} \subseteq \tau(\mathcal{S}) \subseteq \tau'$, by Theorem 2.4$(iii)$, identifying $\tau(\mathcal{S})$ as τ_1, τ' as

τ_2, and $\mathcal{B}$ as $\mathcal{B}_2$, we have $\tau(\mathcal{S}) = \tau'$. In particular, we see that $\mathcal{B}$ is a base for $\tau(\mathcal{S})$. □

2.9 Remark. From Theorem 2.8, it may appear that a base for a topology must be $\cap$-stable. Indeed, by the construction of the weakest topology $\tau(\mathcal{S})$, we moved from $\mathcal{S}$ to $\mathcal{B}(\mathcal{S})$ including all finite intersections of subbase sets, which turned to a base of $\tau(\mathcal{S})$. Moreover, $\mathcal{B}(\mathcal{S})$ was extended to $\tau(\mathcal{S})$ by the mere inclusion in $\tau(\mathcal{S})$ of all unions of base elements, which looks like the fundamental construction of a topology beginning with an almost arbitrary family of subsets.

However, the truth is that, in general, an arbitrary base for a topology need not be $\cap$-stable at all. Firstly, notice that the family of all open balls in a metrizable topology is a base, but it is not $\cap$-stable. In Example 2.9(iii), the collection of all hypercubes in $\mathbb{R}^n$ is a base for the usual topology τ_e^n, but it obviously is not $\cap$-stable either. Secondly, from Problem 2.2, if $\mathcal{B}$ is a base for some τ, the intersection of two base elements if not empty must allow to hold as a subset, another base element, but it does not require that the intersection of these two base sets be a base set.

Now, the system $\mathcal{C}$ of all hypercubes can be extended to an $\cap$-stable family which can be easily guessed, brings us to the set $\mathcal{R}$ of all simple rectangles, which obviously becomes a larger base for τ_e^n compared to $\mathcal{C}$. Thus, $\tau' := \tau(\mathcal{C}) = \tau(\mathcal{R}) \subseteq \tau_e^n$, simply because τ' is the weakest topology containing $\mathcal{C}$ and $\mathcal{R}$, and $\mathcal{C} \subseteq \mathcal{R} \subseteq \tau_e^n$. On the other hand, because $\mathcal{R}$ is a base for τ_e^n and $\tau(\mathcal{R})$, by Theorem 2.4 (ii) the topologies $\tau(\mathcal{C}) = \tau(\mathcal{R})$ and the usual topology τ_e^n must coincide. Recall that metrics d_e and the supremum metric are equivalent on $\mathbb{R}^n$. No wonder that τ_e^n and $\tau(\mathcal{C})$ coincide.

In a nutshell, an idea that every base for a topology is $\cap$-stable is taken out of context from Theorem 2.8. □

2.10 Examples.

(i) In Example 2.3(iii) and Remark 2.9, we argued that the system $\mathcal{C}$ of all hypercubes and the system $\mathcal{R}$ of all simple rectangles are bases for $(\mathbb{R}^n, \tau_e^n)$, of which the second is $\cap$-stable. We can take for a base for $\tau(\mathcal{S})$ in $\mathbb{R}^n$ the system of all simple rectangles with ra-

tional coordinates, which is certainly closed relative to all finite intersections and is also a base for τ_e^n. Because this base is obviously countable, it thus follows that $(\mathbb{R}^n, \tau_e^n)$ is second countable.

(ii) In another scenario of $(\mathbb{R}^n, \tau_e^n)$, the collection of open cylinders of types

$$\pi_i^*((a_i, b_i)) = \mathbb{R} \times \cdots \times \mathbb{R} \times (a_i, b_i) \times \mathbb{R} \times \ \ldots \ \times \mathbb{R},$$

where (a_i, b_i)'s are open bounded intervals in $\mathbb{R}$, $i = 1, \ldots, n$, forms a subbase for τ_e^n. Now, the intersection of two cylinders is a cylinder or the empty set only within the cylinders $\{\pi_i^*((a_i, b_i))\}$ for the same fixed i. Thus, the entire family of cylinders is not $\cap$-stable, but it can be extended to an $\cap$-stable base $\mathcal{B}$ by including in $\mathcal{B}$ the empty set $\emptyset$ and all finite intersections of the subbase rectangles. Base $\mathcal{B}$ evidently contains $\mathcal{R}$ and is also a base for $\tau(\mathcal{R})$ and τ_e^n. □

PROBLEMS

2.1 Let (X, τ) be a topological space and let $\mathcal{B} \subseteq \tau$. Show that $\mathcal{B}$ is a base for τ if and only if for every open set $O \in \tau$ and each point $x \in O$, there is a subset B of O such that $x \in B \in \mathcal{B}$.

2.2 Show that $\mathcal{B} \subseteq \mathcal{P}(X)$ is a base for a topology on X if and only if two conditions are met:

(i) each $x \in X$ belongs to at least one set $B \in \mathcal{B}$ (or equivalently, $X = \bigcup_{B \in \mathcal{B}} B$).

(ii) $\forall B_1, B_2 \in \mathcal{B}$ and $\forall x \in B_1 \cap B_2$, $\exists B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \cap B_2$.

[*Hint*: Use the following steps.

1) If $\mathcal{B}$ is a base, then apply Theorem 2.2(ii).
2) Let $\tau = \left\{ \bigcup_{B \in \mathcal{B}'} B : \forall\, \mathcal{B}' \subseteq \mathcal{B} \right\}$. Show that τ is a topology on X and that $\mathcal{B}$ is a base for τ.]

2.3 Let $\mathcal{B}$ be a base for a topology τ on X. In as much as $\mathcal{B}$, in particular, is a subbase on X, it also generates the weakest topology $\tau(\mathcal{B})$ and hence $\tau(\mathcal{B}) \subseteq \tau$. Is $\tau(\mathcal{B}) = \tau$?

2.4 Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be bases for the topologies (X, τ_1) and (X, τ_2), respectively. Prove the equivalence of the following statments.

(i) For each $x \in X$ and each base set $A_x \in \mathcal{B}_1$ that contains x, there is a base set $B_x \in \mathcal{B}_2$ that contains x and $B_x \subseteq A_x$.

(ii) τ_2 is finer than τ_1.

2.5 Let τ_l denote the topology on the real line generated by all half-open intervals of type $[a, b)$ where $a, b \in \mathbb{R}$. This topology is called the *lower limit topology*. Show that $\{[a, b)\colon a, b \in \mathbb{R}\}$ is a base for τ_l and that τ_l is strictly finer than τ_e, the usual topology on the real line.

2.6 Show that the collection of all sets on the real line of types (a, ∞) and $(-\infty, b)$ is a subbase for the usual topology $(\mathbb{R}, \tau_e)$.

2.7 Show that any base and subbase rectangles in Example 2.3(iii) and Example 2.10(ii), respectively, are open sets.

2.8 Show that the usual topology τ_e on $\mathbb{R}^n$ and the product topology τ_P generated by $\underset{i=1}{\overset{n}{\times}} \tau_e^{(i)}$ coincide.

NEW TERMS:

base for a topology, 188
base sets, 188
base for a topology criterion, 188, 190, 194
open rectangle (parallelepiped), 190
open parallelepiped (rectangle), 190
base (simple) rectangle, 190
simple (base) rectangle, 190
subbase, 191, 192
topology generated by a subbase, 191
open cylinders, 194
lower limit topology, 195

3. HAUSDORFF TOPOLOGICAL SPACES

We start this section with the construction of a rudimentary product topology that will be revisited in Section 5 and which presently plays a provisional role in connection with the uniqueness of limits in some topological spaces. The entire section is dedicated to the convergence of sequences, first introduced in Definition 1.10. We also acquaint the reader with one of the most common classes of topological spaces, the Hausdorff space, in which all convergent sequences have unique limits, a very useful feature we come to appreciate in metric spaces.

3.1 Definition. Let (X_1, τ_1) and (X_2, τ_2) be two topological spaces. Define

$$\tau_1 \times \tau_2 = \{A \times B : A \in \tau_1, B \in \tau_2\}.$$

Then, obviously, $\bigcup_{U \in \tau_1 \times \tau_2} U = X_1 \times X_2$ and, in addition, by Problem 5.6(a) of Chapter 1, $\tau_1 \times \tau_2$ is $\cap$-stable. Therefore, by Theorem 2.8, we can form a topology $\tau(\tau_1 \times \tau_2)$ generated by $\tau_1 \times \tau_2$ (as a base), denote it by τ_P, and call it the *product topology*. (We discuss more general variants of τ_P in Section 5.) □

We continue with a useful lemma needed in the sequel.

3.2 Lemma. *A sequence $\{\xi_n = (x_n, y_n)\} \subseteq X_1 \times X_2$ is convergent to a point $\xi = (x, y)$ if and only if* □*marginal sequences*□ *$\{x_n\}$ and $\{y_n\}$ converge to x and y, respectively*.

Proof. (i) Let $\xi_n \to \xi$ in τ_P. Then the marginal sequences $\{x_n\}$ and $\{y_n\}$ converge to x and y, respectively. Indeed, let O_x and O_y be open neighborhoods of x and y. Then, $O_x \times O_y$ is a base set in τ_P and an open neighborhood of ξ. Because $\{\xi_n\}$ is convergent, there is an $N(O_x \times O_y)$-tail of $\{\xi_n\}$. Hence, marginal tails $\{x_N, x_{N+1}, \dots\}$ and $\{y_N, y_{N+1}, \dots\}$ are subsets of O_x and O_y, respectively. Therefore, $x_n \to x$ and $y_n \to y$.

(ii) Conversely, let $\{x_n\}$ and $\{y_n\}$ be two sequences from X_1 and X_2 that converge to x and y, respectively, and let W_ξ be an open neighborhood of $\xi = (x, y)$. Then, by the same arguments as in part (i), there is an open subset (as a base set) $O_1 \times O_2 \subseteq W_\xi$ such that $\xi \in O_1 \times O_2$; that is, O_1 and O_2 are open neighborhoods of x and y. Thus, there is an $M(O_1)$-tail of $\{x_n\}$ and $N(O_2)$-tail of $\{y_n\}$ implying that the $R(O_1 \times O_2)$-tail (where $R = \max\{M, N\}$) is a subset of $O_1 \times O_2$ and thus of W_ξ. □

Now we prove a simple but important statement that if τ_1 and τ_2 are first countable, then so is τ_P.

3.3 Proposition. *Let (X_1, τ_1) and (X_2, τ_2) be first countable topological spaces. Then $(X_1 \times X_2, \tau_P)$ is also a first countable topological space.*

Proof. Let $\xi = (x, y) \in X_1 \times X_2$. Because X_1 and X_2 are first countable, there are countable neighborhood bases $\mathcal{B}_x = \{U_x^j : j = 1, 2, \dots\}$ and $\mathcal{B}_y = \{V_y^k : k = 1, 2, \dots\}$ of x and y, respectively. Define $\mathcal{B}_\xi = \{U_x^j \times V_y^k : j, k = 1, 2, \dots\}$ and show that it is a neighborhood base for ξ. Without loss of generality and to simplify the proof we assume that all neighborhoods involved are open.

Now, let W_ξ be an open neighborhood of ξ. Then, since $\tau_1 \times \tau_2$ is a base for τ_P, W_ξ can be represented as $W_\xi = \bigcup_{\text{some } A \in \tau_1, B \in \tau_2} (A \times B)$. Hence, there is an $A \times B \subseteq W_\xi$ such that $\xi \in A \times B$ or $x \in A$ and $y \in B$. Because $\mathcal{B}_x$ and $\mathcal{B}_y$ are neighborhood bases for x and y and A and B are open neighborhoods of x and y, there are U_x^j and V_y^k that are subsets of A and B. Consequently, $U_x^j \times V_y^k \subseteq A \times B \subseteq W_\xi$. Thus, $\mathcal{B}_\xi$ is indeed a neighborhood base for ξ and it is countable. □

In various applications encountered in topological spaces we deal with (including those in forthcoming Chapters 5 and 6), every convergent sequence has exactly one limit. An important representative of this class is introduced in the definition below.

3.4 Definition. A topological space (X, τ) is said to be *Hausdorff* (or *separated* or T_2) if every two distinct points, $x, y \in X$, possess disjoint neighborhoods. □

T_2 is a designation of the *second separation axiom*. (The character T stands for the German word *Trennungsaxiom*.) Other separation axioms are introduced and discussed in [Dsh2]. As mentioned, the next proposition (which is hardly a challenge for the reader) is a consequence of Hausdorff spaces.

In Example 1.11(i) we saw that a sequence in general can have more than one limit (speaking conservatively). The "Hausdorfness" of a topological space is a noteworthy and often used restriction emulating metric spaces.

3.5 Proposition. *Let* (X, τ) *be a Hausdorff topological space, and let* $\lim_{n\to\infty} x_n = x$, *and* $\lim_{n\to\infty} x_n = y$. *Then* $x = y$. □

(See Problem 3.1.)

3.6 Example. Let (X, d) be a metric space and let $(X, \tau(d))$ be the corresponding metrizable topological space. With x_1 and x_2 being distinct points of X, we pick two open balls, $B(x_1, r)$ and $B(x_2, r)$, with $r = \frac{1}{3}d(x_1, x_2)$. It follows that the balls are neighborhoods of x_1 and x_2, respectively, and that $B(x_1, r) \cap B(x_2, r) = \emptyset$. This immediately implies that any metrizable topological space is Hausdorff. □

3.7 Remark. We want to raise an important question: when can Proposition 3.5 be reversed, that is, when does the uniqueness of limits imply that the space (X, τ) is Hausdorff? In Example 1.16(ii) we argued that in a cocountable topological space, the only sequences with at most finitely many distinct terms can have limits and they are unique, yet this space is not Hausdorff. Indeed, any two open neighborhoods O_x and O_y of any two points x and y are not disjoint, because they differ only upon at most countable sets. So, what will then a reasonable sufficient condition be that can make a topological space with the "uniqueness of limits property" Hausdorff? □

We begin with the lemma below that builds promised bridges between uniqueness of limits of sequences and Hausdorffness of

spaces. The same result is generalized and applied to filters and nets in [Dsh2].

3.8 Lemma. *Let (X, τ) be a topological space. Then (X, τ) is Hausdorff if and only if the diagonal $D = \{(x, x) \in X^2\}$ is closed in the product topology (X^2, τ_P).*

Proof.

(i) Let (X, τ) be Hausdorff and let $\xi = (x, y) \in D^c$. Thus, $x \neq y$ and due to the "Hausdorfness" of X, x and y have disjoint open neighborhoods O_x and O_y. In as much as they are disjoint, obviously $O_x \cap O_y$ is disjoint with D. Furthermore, because $O_x \cap O_y$ is a base set of τ_P, it is open and thus ξ is an interior point of $D^c \Rightarrow D$ is closed.

(ii) Conversely, if D^c is open, it can be represented as a union of base open sets, that is, as $D^c = \bigcup_{\text{some } A,B \in \tau} A \times B$. Then given a point $(x, y) \in D^c$, at least one of these sets $A \times B$, say $O_x \times O_y \subseteq D^c$, must contain (x, y). Thus, $x \in O_x$ and $y \in O_y$ and, consequently, O_x and O_y are open neighborhoods of x and y, respectively. They are disjoint, because $O_x \times O_y \subseteq D^c$. Hence, (X, τ) is Hausdorff. □

Now we are ready for the main result of this section.

3.9 Theorem. *Let (X, τ) be a first countable topological space in which limits of all convergent sequences are unique. Then, (X, τ) is Hausdorff.*

Proof. Let D be the diagonal in the product topology τ_P and let $\xi \in \bar{D}$. By Proposition 3.3, τ_P is first countable. Therefore, by Proposition 1.14, there must be a sequence $\{(x_n, x_n)\} \subseteq D$ such that $(x_n, x_n) \to \xi = (x, y)$. By Lemma 3.2, the marginal sequence $\{x_n\}$ must converge to x and y. However, since by our assumption, all convergent sequences in X must have unique limits, $x = y$ and hence $\xi \in D$ implying that $\bar{D} \subseteq D$ and thus D is closed. By Lemma 3.8, (X, τ) is Hausdorff. □

PROBLEMS

3.1 Prove Proposition 3.5. *Let* (X, τ) *be a Hausdorff topological space,* $\lim_{n\to\infty} x_n = x$, *and let* $\lim_{n\to\infty} x_n = y$. *Then* $x = y$.

3.2 Show that any one-point set in a Hausdorff space is closed.

3.3 Show that any metric space is first countable.

3.4 Prove that any separable metric space is second countable.

3.5 Is it true that any first countable topological space is also second countable?

3.6 Prove that if a topological space is second countable, then it is separable and first countable.

3.7 Show that subspaces of second countable spaces are second countable.

3.8 Prove that the usual topology τ_e^n on $\mathbb{R}^n$ and the product topology τ_P on $\mathbb{R}^n$ generated by $\underset{i=1}{\overset{n}{\times}} \tau_e^{(i)}$ coincide, where $\tau_e^{(i)} = \tau_e$ is the usual topology on $\mathbb{R}$.

NEW TERMS:

product topology (of two spaces), 197
marginal sequences, 197
Hausdorff (separated, T_2) topological space 199
separated topological space, 199
T_2 space, 199
second separation axiom, 199
Trennungsaxiom, 199
diagonal, 200
Hausdorff space, criterion of, 200

4. CONTINUITY IN TOPOLOGICAL SPACES

Except for a brief introduction of sequences (being restricted versions of functions) in Section 3, in the present section functions appear for the first time in conjunction with topologies. Naturally, their most significant quality that we look into is continuity.

After a first acquaintance with continuity in metric spaces (Section 4, Chapter 2), the reader is now well prepared to its surprising generalization for topological spaces and a striking similarity between Theorem 4.2 below and Theorem 4.3, Chapter 2, with respect to a key continuity criterion. Again, we observe some other continuity properties, typical for metric spaces and holding for special topological spaces, yet more general than those in metric spaces. One of them deals with an important relationship between convergence of sequences and continuity of functions initiated in Chapter 2 (earlier formulated as Theorem 4.4 and pledged to be proved in this section).

4.1 Definitions.

(i) A function $f\colon (X,\tau) \to (Y,\tau_1)$ is said to be *continuous* at a point $a \in X$ if, for every neighborhood $W_{f(a)}$, there is a neighborhood U_a such that $f_*(U_a) \subseteq W_{f(a)}$.

This is equivalent to the following definition. *f is continuous at a, if for every neighborhood $W_{f(a)}$ of $f(a)$, $f^*(W_{f(a)})$ is a neighborhood of a* (see Problem 4.1).

(ii) The function f is said to be *continuous on* X (or simply *continuous*) if it is continuous at each point $a \in X$. □

4.2 Theorem. *Let $f : (X,\tau) \to (Y,\tau_1)$ be a function. Then the following are equivalent.*

(i) *f is continuous.*

(ii) *The inverse image under f of any open set $H \in \tau_1$ is open, that is, is an element of τ.*

Proof.

$(i) \Rightarrow (ii)$. Let $H \in \tau_1$. For each point $a \in f^*(H)$, $f(a) \in H$ and therefore $f(a)$ is an interior point of H. Specifically, H is a neighborhood of $f(a)$. Because f is continuous at a, there is a neighborhood U_a such that $f(U_a) \subseteq H$. Because the inclusion is preserved under the inverse, we have

$$U_a \subseteq f^*(f(U_a)) \subseteq f^*(H),$$

which implies that $f^*(H)$ contains a neighborhood for each of its points. Hence, $f^*(H)$ is itself a neighborhood for all of its points. Therefore, by Proposition 1.6, $f^*(H)$ is open, i.e. is an element of τ.

$(ii) \Rightarrow (i)$. Let $a \in X$ and let $W_{f(a)}$ be a neighborhood of $f(a)$. Then, there exists an open set $H \in \tau_1$ such that $f(a) \in H \subseteq W_{f(a)}$. By assumption (ii), $f^*(H)$ an element of τ. Because obviously $a \in f^*(H)$, $f^*(H)$ is a neighborhood of a and thus $f^*(W_{f(a)})$ is also a neighborhood of a. Consequently, we have continuity of f at a. □

Let (X, τ) be a topological space. Recall that τ^c denotes the collection of all closed sets.

4.3 Proposition. *A function $f : (X, \tau) \to (Y, \tau_1)$ is continuous on X if and only if the inverse image under f of any closed set $O^c \in \tau_1^c$ is closed in (X, τ).* □

(See Problem 4.2.)

4.4 Proposition. *Let (X, τ), (Y, τ_1), and (Z, τ_2) be topological spaces and let $f : X \to Y$ and $g : Y \to Z$ be continuous functions. Then the function $g \circ f : X \to Z$ is continuous.* □

(See Problem 4.3.)

4.5 Definition. Let (X, τ) be a topological space and let $[X, Y, f]$ be a function. Define

$$\tau_Q = \{B \subseteq Y : f^*(B) \in \tau\},$$

i.e., $f^{**}(\tau_Q) \subseteq \tau$. By the following arguments (Remarks 4.6), τ_Q is a topology and it contains any topology relative to which f is continuous. τ_Q is called the *quotient topology induced on Y by f*. □

[Recall that f^* is defined on $\mathcal{P}(X)$; consequently, f^{**} is a function acting on $\mathcal{P}(\mathcal{P}(X))$. See Definition 2.1$(vii)$ and $(viii)$, Chapter 1.]

4.6 Remarks.

(i) τ_Q is indeed a topology:

1) $\emptyset,\ Y \in \tau_Q$.

2) $B_1, \ldots, B_n \in \tau_Q \ \Rightarrow\ f^*(\bigcap_{k=1}^{n} B_k) = \bigcap_{k=1}^{n} f^*(B_k) \in \tau$ (as the intersection of open sets) $\Rightarrow\ \bigcap_{k=1}^{n} B_k \in \tau_Q$.

3) A similar consideration can be used to show that τ_Q contains all unions.

(ii) τ_Q is the largest topology on Y relative to which f is continuous. This follows directly from Definition 4.5. □

4.7 Example. Let $X = \mathbb{R},\ \tau = \{\mathbb{R},\ \emptyset,\ (-1, 2],\ [0, 3),\ [0, 2],\ (-1, 3),\ (-1, 1)\}$ and let $f(x) = x^2$ defined as $f: \mathbb{R} \to \mathbb{R} = Y$. It is clear that $\mathbb{R}$, $\emptyset$ and $[0, 1)$ are the only subsets of Y whose inverse images are in τ. Therefore, $\{\mathbb{R}, \emptyset, [0, 1)\}$ is the quotient topology on Y. □

By Theorem 4.2, $f: (X, \tau) \to (Y, \tau')$ is continuous if and only if $f^{**}(\tau') \subseteq \tau$. However, if we know a generator $\mathcal{S}$ of τ', then condition (ii) of Theorem 4.2 can be relaxed as per the following theorem.

4.8 Theorem. *Let $f: (X, \tau) \to (Y, \tau(\mathcal{S}))$ [where $\tau(\mathcal{S})$ is the topology generated by a subbase $\mathcal{S}$]. Then f is continuous if and only if $f^{**}(\mathcal{S}) \subseteq \tau$.*

Proof. If f is continuous, then, in particular, $f^{**}(\mathcal{S}) \subseteq \tau$. Assume that $f^{**}(\mathcal{S}) \subseteq \tau$ and introduce the quotient topology τ_Q induced by f. Thus, $\mathcal{S} \subseteq \tau_Q$, which implies that $\tau(\mathcal{S}) \subseteq \tau_Q$, for $\tau(\mathcal{S})$

is the smallest topology containing $\mathcal{S}$. Then in as much as $f^{**}(\tau_Q) \subseteq \tau$, we have

$$f^{**}(\tau(\mathcal{S})) \subseteq f^{**}(\tau_Q) \subseteq \tau. \qquad \square$$

4.9 Theorem. *Let $f : (X, \tau) \to (Y, \tau')$ be a map continuous at some point $x \in X$. If $\{x_n\}$ is a sequence convergent to x, the sequence $\{f(x_n)\}$ is convergent to $f(x)$.* $\square$

(See Problem 4.10.)

4.10 Remark. Let (X_1, τ_1) and (X_2, τ_2) be two topological spaces and τ_P be the product topology on X^2. Obviously both projection maps π_1 and π_2 are continuous, because $\pi_1^*(O_1) = O_1 \times X_2 \in \tau_1 \times \tau_2 \subseteq \tau_P$. So, $\pi_1^{**}(\tau_1) \subseteq \tau_P$. Thus, π_1 is continuous and so is π_2. It offers an alternative argument to part (i) in Lemma 3.2 that a sequence $\{\xi_n = (x_n, y_n)\} \subseteq X_1 \times X_2$ is convergent to a point $\xi = (x, y)$, then marginal sequences $\{x_n\}$ and $\{y_n\}$ converge to x and y, respectively.

Indeed, by Theorem 4.9, the marginal sequences $\{\pi_1(\xi_n) = x_n\}$ and $\{\pi_2(\xi_n) = y_n\}$ converge. More specifically, if $(x_n, y_n) \to (x, y)$, then $x_n \to x$ and $y_n \to y$. $\square$

Theorems 4.8 and 4.9 and the next theorem form an analogue to Heine's Theorem 4.4, Chapter 2, which was valid for metric spaces. The statement in Theorem 4.9 has no restriction as to the nature of topological spaces (X, τ) and (Y, τ'), although its converse needs to be constrained by the condition that (X, τ) is first countable.

4.11 Theorem. *Let $f: (X, \tau) \to (Y, \tau')$ be a map and let (X, τ) be first countable. If for any sequence $\{x_n\}$ convergent to a point $x \in X$, the sequence $\{f(x_n)\}$ converges to $f(x)$, then f is continuous at x.*

Proof. To prove this theorem, we assume that f is not continuous at x; then select a sequence $\{x_n\}$ convergent to x such that $\{f(x_n)\}$ does not converge to $f(x)$. The assumption that (X, τ) is first countable is essential in the selection of a convergent sequence $\{x_n\}$, which otherwise need not exist. If f is not continuous at x, there is a neighborhood $W_{f(x)}$ such that $f^*(W_{f(x)})$ is not a neighbor-

hood of x, or equivalently, there is no neighborhood U_x such that $f(U_x) \subseteq W_{f(x)}$. [Otherwise, if $f(U_x) \subseteq W_{f(x)}$, then

$$U_x \subseteq f^*(f_*(U_x)) \subseteq f^*(W_{f(x)}).$$

This would contradict our assumption. (See Figure 4.1.)]

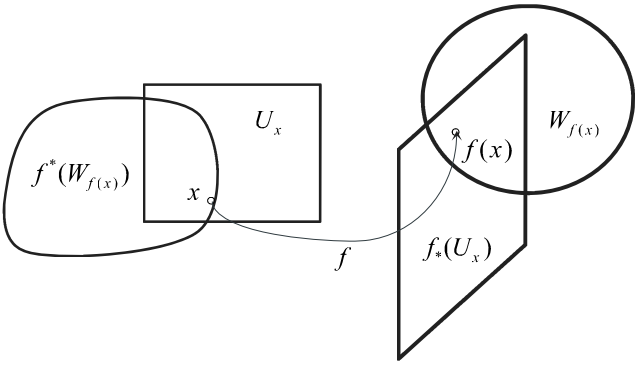


Figure 4.1

Specifically, it follows that, for each base neighborhood $B \in \mathcal{B}_x$, $f_*(B)$ is not a subset of $W_{f(x)}$. Because (X, τ) is first countable, there is a countable neighborhood base $\mathcal{B}_x = \{B_1, B_2, \dots\}$ which can always be assumed to be monotone nonincreasing. Now, each B_i contains at least one point, say x_i, such that $f(x_i) \notin W_{f(x)}$, which immediately yields that the sequence $\{f(x_n)\}$ is not in $W_{f(x_1)}$ and, thus, does not converge to $f(x)$. However, by Proposition 1.13, $x_n \to x$. □

Theorem 4.11 leads to some useful applications. Prior to that we need the following.

4.12 Definition. Let $f, g : X \to Y$ be two functions. Define $f \otimes g : X^2 \to Y^2$ as $(f \otimes g)(x_1, x_2) = (f(x_1), g(x_2))$. Consequently, we have $[\mathcal{P}(X^2), \mathcal{P}(Y^2), (f \otimes g)_*]$ specified as

$$(f \otimes g)_*(A) = \{(f(x_1), g(x_2)) : (x_1, x_2) \in A\},\ A \subseteq X^2$$

and $[\mathcal{P}(Y^2), \mathcal{P}(X^2), (f \otimes g)^*]$ as

$$(f \otimes g)^*(B) = \{(f^*\{y_1\}, g^*\{y_2\}) : (y_1, y_2) \in B\}, B \subseteq Y^2.$$

In particular, it is readily seen that

$$(f \otimes g)_*(A_1 \times A_2) = f_*(A_1) \times g_*(A_2) \tag{4.12}$$

and

$$(f \otimes g)^*(B_1 \times B_2) = f^*(B_1) \times g^*(B_2). \tag{4.12a}$$

□

4.13 Lemma. *Let $f, g\colon (X, \tau) \to (Y, \tau')$ be two continuous maps. If (Y, τ') is Hausdorff, then the set $S = \{x \in X : f(x) = g(x)\}$ is closed in (X, τ).*

Proof. Because f and g are continuous, the map $f \otimes g : X^2 \to Y^2$ is continuous relative to the respective product topologies. Indeed, by Theorem 4.8, $f \otimes g$ is continuous if and only if $(f \otimes g)^{**}(\tau' \times \tau') \subseteq \tau_P$, where τ_P is the product topology on X^2. Since $f^*(B_1)$ and $g^*(B_2)$ are open in X, when B_1 and B_2 are open in Y, it follows from (4.12a) that $(f \otimes g)^*(B_1 \times B_2) \in \tau \times \tau \subseteq \tau_P$.

In as much as by the assumption (Y, τ') is Hausdorff, by Lemma 3.8, the diagonal D in $Y \times Y$ is closed. Hence, the set S, as the inverse image of the diagonal D under the continuous map $f \otimes g$ must be closed, as per Proposition 4.3. □

4.14 Proposition. *Let $f, g\colon (X, \tau) \to (Y, \tau')$ be two continuous maps that coincide on some dense set in X. If (X, τ) is first countable and if (Y, τ') is Hausdorff, then $f = g$ on X.* □

Thus, it follows that a continuous function is well-defined on a dense set. The proof of this proposition is the subject of Problem 4.11.

4.15 Example. If $f, g\colon (\mathbb{R}^n, \tau_e) \to (\mathbb{R}^n, \tau_e)$ are continuous maps that coincide on the set $\mathbb{Q}^n$ of all vectors with rational coordinates, then f and g are identical on $\mathbb{R}^n$. This fact takes into account that $(\mathbb{R}^n, \tau_e)$ is Hausdorff and first countable. □

4.16 Definition. Let (X, τ) and (Y, τ') be two topological spaces. A bijective map $[X, Y, f]$ is called a *homeomorphism* if both f and f^{-1} are continuous. The topological spaces (X, τ) and (Y, τ') are then called *homeomorphic.* We write $X \sim Y$. If f fails to be surjective, then f is called an *embedding of X into Y*. X is also said to be *embedded in Y by f*. □

4.17 Remark. It is not difficult to see that the homeomorphic property applied to a family of topological spaces between any fixed carriers X and Y defines an equivalence relation (show it; Problem 4.12). □

PROBLEMS

4.1 Show that f is continuous at a point a if and only if for every neighborhood $W_{f(a)}$, $f^*(W_{f(a)})$ is a neighborhood of a.

4.2 Prove Proposition 4.3. *A function $f: (X, \tau) \to (Y, \tau_1)$ is continuous on X if and only if the inverse image under f of any closed set $O^c \in \tau_1^c$ is closed in (X, τ).*

4.3 Prove Proposition 4.4. *Let (X, τ), (Y, τ_1), and (Z, τ_2) be topological spaces and let $f: X \to Y$ and $g: Y \to Z$ be continuous functions. Then the function $g \circ f: X \to Z$ is continuous.*

4.4 Let $f: (X, \tau) \to (Y, \tau_1)$ be a function such that $f(x) = x$, $X = Y = \mathbb{R}$, $\tau = \{\mathbb{R}, \emptyset, \{1\}, [1, 3)\}$, and $\tau_1 = \{\mathbb{R}, \emptyset, \{2\}, [2, 4)\}$. Is f continuous?

4.5 Under the conditions of Problem 4.4, set $f(x) = x + 1$. Is f continuous?

4.6 Let $f: (X, \tau) \to (Y, \tau')$ be a map. Show that f is continuous at a point $x \in X$ if and only if, for any base neighborhood $B_{f(x)}$ of the point $f(x)$, $f^*(B_{f(x)})$ is a neighborhood of x.

4.7 Under the condition of Problem 4.6, assume that $\tau' = \tau(d)$, that is, (Y, τ') is a metrizable topological space.

a) Show that f is continuous at $x \in X$ if and only if the inverse image under f of any open ball $B_d(f(x), \varepsilon)$ is a neighborhood of x.

b) Show that, for each open ball $B_d(f(x), \varepsilon)$ there is a neighborhood $U_x(\varepsilon)$ such that

$$f(U_x(\varepsilon)) \subseteq B_d(f(x), \varepsilon)\,.$$

4.8 Let $f\colon (X, \tau) \to (Y, \|\cdot\|)$ be a map, where Y is an NVS over a field $\mathbb{F}$. Show that f is continuous at $x \in X$ if and only if, for every $\varepsilon > 0$, there is a neighborhood $U_x(\varepsilon) \in \mathcal{U}_x$ such that for each $y \in U_x(\varepsilon)$, $\| f(x) - f(y) \| < \varepsilon$.

4.9 Prove the following statement. *Let $f\colon (X, \tau) \to (\mathbb{R}^n, \|\cdot\|_e)$, where (X, τ) is a topological space. Then f is continuous at a point $x \in X$ if and only if, for every $\varepsilon > 0$, there is a neighborhood $U_x(\varepsilon) \in \mathcal{U}_x$ such that, for all points $y \in U_x(\varepsilon)$, $\|f(x) - f(y)\|_e < \varepsilon$.*

4.10 Prove Theorem 4.9. *Let $f : (X, \tau) \to (Y, \tau')$ be a map continuous at some point $x \in X$. If $\{x_n\}$ is a sequence convergent to x, the sequence $\{f(x_n)\}$ is convergent to $f(x)$.*

4.11 Prove Proposition 4.14. *Let $f, g\colon (X, \tau) \to (Y, \tau')$ be two continuous maps that coincide on some dense set in X. If (X, τ) is first countable and if (Y, τ') is Hausdorff, then $f = g$ on X.*

4.12 Prove the statement posed in Remark 4.17. *Show that the homeomorphic property applied to a family of topological spaces between any fixed carriers X and Y defines an equivalence relation.*

4.13 Show that $(\mathbb{R}, \tau_e)$ is homeomorphic to $(-1, 1)$ with the corresponding relative topology on $(-1, 1)$.

4.14 Is $(\mathbb{R}, \tau_e)$ homeomorphic to $[-1, 1]$?

NEW TERMS:

function continuous at a point, 203
function continuous at a point, criterion of, 203
continuous function, 203
continuity of a function, criterion of, 203, 204, 205, 206
composition of continuous functions, 204
quotient topology, 205
continuity of projection maps, 206
continuous function on a dense set, 208, 210
homeomorphism, 209
homeomorphic topological spaces, 209
embedding, 209
embedded set, 209

5. PRODUCT TOPOLOGY

Let $(Y_1, \tau_1), \ldots, (Y_n, \tau_n)$ be "factor" topological spaces. One of the reasonable ways to define a topology on the Cartesian product $\Omega = Y_1 \times \cdots \times Y_n$ (earlier introduced in Section 3 for two spaces) is to take the collection

$$\mathcal{B} = \overset{n}{\underset{i=1}{\times}} \tau_i : = \tau_1 \times \tau_2 \times \cdots \times \tau_n$$

$$= \left\{ \prod_{i=1}^{n} O_i : O_i \in \tau_i,\ i = 1, \ldots, n \right\},$$

for a family of open rectangles. Because $\mathcal{B}$ is $\cap$-stable we use it as a base for the topology it generates. Indeed, by Theorem 2.8, $\mathcal{B}$ is a base for $\tau(\mathcal{B})$ that includes all unions of elements of $\mathcal{B}$. We call $\tau(\mathcal{B})$ the *product topology on* Ω and denote it by τ_P.

The proposition below allows a significant reduction of the base $\mathcal{B} = \overset{n}{\underset{i=1}{\times}} \tau_i$ for τ_P provided we have the knowledge of bases for the respective factor topologies.

5.1 Proposition. *Let* $\widetilde{\mathcal{B}} = \overset{n}{\underset{i=1}{\times}} \mathcal{B}_i = \mathcal{B}_1 \times \cdots \times \mathcal{B}_n$

$$= \left\{ \prod_{i=1}^{n} B_i \ : B_i \in \mathcal{B}_i,\ i = 1, \ldots, n \right\},$$

where $\mathcal{B}_i$ *is a base for ith factor topology* $\tau_i,\ i = 1, \ldots, n$. *Then* $\widetilde{\mathcal{B}}$ *is also a base for* τ_P. □

The reader can readily handle the proof in Problem 5.8.

Any element of $\widetilde{\mathcal{B}}$ is called a *simple* or *base rectangle.*

The construction below is an effort to reduce further a generator for the product topology on finitely many spaces.

5.2 Proposition. *Let* $\mathcal{S} = \overset{n}{\underset{i=1}{\times}} \mathcal{S}_i$ *defined as*

$$\mathcal{S} = \left\{ \prod_{i=1}^{n} S_i : S_i \in \mathcal{S}_i,\ i = 1, \ldots, n \right\},$$

where $\mathcal{S}_i$ is a subbase for $\tau_i = \tau(\mathcal{S}_i),\ i = 1, \ldots, n$. Then, $\mathcal{S}$ is a subbase for τ_P, that is, $\tau(\mathcal{S}) = \tau_P$. □

Notice that the base for τ_P of Proposition 5.1 involved in the construction of τ_P contains finite intersections of subbase elements.

Any element of $\mathcal{S}$ in Proposition 5.2 is called a *subbase rectangle.* (The proof of Proposition 5.2 is assigned to Problem 5.9.)

5.3 Proposition. *Let $\mathcal{S}'$ be the set of all unit subbase cylinders*

$$\mathcal{S}' = \bigcup_{i=1}^{n} \pi_i^{**}(\mathcal{S}_i) = \{\pi_i^*(S_i) : S_i \in \mathcal{S}_i,\ i = 1, \ldots, n\},$$

where $\mathcal{S}_i$ is a subbase for τ_i. Then $\mathcal{S}'$ is a subbase for $\tau_P = \tau(\mathcal{S})$. □

(See Problem 5.10.)

5.4 Example. According to Problem 3.8, the usual topology τ_e^n on $\mathbb{R}^n$ and the product topology τ_P on $\mathbb{R}^n$ generated by $\overset{n}{\underset{i=1}{\times}} \tau_e^{(i)}$ coincide, where $\tau_e^{(i)} = \tau_e$ is the usual topology on $\mathbb{R}$. By Proposition 5.1, the usual topology τ_e^n on $\mathbb{R}^n$ is thus generated by simple rectangles of the form $\prod_{i=1}^{n} (a_i, b_i)$, where $(a_i, b_i) \subseteq \mathbb{R}$ are base elements of $(\mathbb{R}, \tau_e)$. In particular, the system of all rational rectangles is also a base for $\tau_P = \tau_e^n$. According to Proposition 5.3, the system $\mathcal{S}'$ of all unit cylinders $\{\pi_i^*((a_i, b_i)) : (a_i, b_i) \subseteq \mathbb{R},\ i = 1, \ldots, n\}$ is a subbase for τ_e^n. [See also Example 2.10(ii).] Once they intersect, for different indices i, they form simple cylinders (for some i's) or rectangles if the intersection is of the type $\bigcap_{i=1}^{n} \pi_i^*((a_i, b_i))$ whose family becomes a base for τ_e^n. Therefore, we can say that the usual topology on $\mathbb{R}^n$ is generated by the set of all simple cylinders. □

From Remark 4.10 we learned that the projection maps are continuous relative to the product topology. Furthermore, we have the following.

5.5 Theorem. *Let* $\Omega = \prod_{i=1}^{n} Y_i$ *and* $\pi_i : \Omega \to Y_i$ *be the* i*th projection map,* $i = 1, \ldots, n$*. Then, the product topology* τ_P *on* Ω *is the weakest topology with respect to which each projection is continuous.*

Proof. Let τ be a topology on Ω, relative to which each projection is continuous, that is, suppose $\pi_i^{**}(\tau_i) \subseteq \tau$. Then, $\mathcal{B} = \overset{n}{\underset{i=1}{\times}} \tau_i \subseteq \tau$. Indeed, if $O \in \overset{n}{\underset{i=1}{\times}} \tau_i$, then $O = \bigcap_{i=1}^{n} \pi_i^*(O_i)$ (where $O_i \in \tau_i$) with $\pi_i^*(O_i) \in \tau$, and because τ is $\cap$-stable, $O \in \tau$. Finally, because $\mathcal{B}$ is a base for τ_P, by Theorem 2.4 (i), $\tau_P \subseteq \tau$. □

5.6 Remark. Now, we extend the notion of the product topology of finitely many factor spaces to the product topology on the Cartesian product of arbitrary many factor spaces. We therefore assume that $\{(Y_x, \tau_x) : x \in X\}$ is an arbitrary indexed family of topological spaces. Let us consider two different models of product topologies on the Cartesian product $\Omega = \prod_{x \in X} Y_x$. One of them, called the *box topology* (in notation τ_B), is similar to that for finitely many factor spaces and is subject to the following construction. As a base for τ_B, we take the system of *box rectangles*,

$$\mathcal{B}_B = \left\{ \prod_{x \in X} O_x : O_x \in \tau_x \right\}, \tag{5.6}$$

or even a weaker base,

$$\mathcal{B}_B = \left\{ \prod_{x \in X} B_x : B_x \in \mathcal{B}_x \right\}. \tag{5.6a}$$

Hence, this way the formed box topology τ_B is not different from its version for finitely many factor spaces and it likewise preserves the continuity of projection maps. The box topology was introduced by the Austrian mathematician Heinrich Tietze in his 1923 paper "Beiträge zur allgemeinen Topologie I" (see [Tiet]).

There is another, "more economical" topology on Ω, which also preserves continuity of projection maps, and in addition, leads to a tamer formation of the widely used "pointwise topology" (which the box topology does not as we show it).

5.7 Definition. Let us define the topology τ_P on Ω through the base

$$\mathcal{B} = \left\{ \pi_A^* \left(\prod_{x\in A} O_x \right) : A\text{'s are finite subsets of } X,\ O_x \in \tau_x \right\}, \tag{5.7}$$

where

$$\left[\prod_{x\in X} Y_x \,,\ \prod_{\alpha\in A} Y_\alpha \,,\ \pi_A \right]$$

is the A-projection map, according to Definition 5.4 (v), Chapter 1. In other words,

$$\mathcal{B} = \left\{ \prod_{x\in X} O_x : O_x \in \tau_x \right\}, \tag{5.7a}$$

where $O_x = Y_x$, except for finitely many indices $x \in X$. Thus, all elements of $\mathcal{B}$ are simple cylinders (see Definition 5.4 (iii), Chapter 1). The topology τ_P (or in notation, rather $\left(\prod_{x\in X} Y_x, \tau_P \right)$) generated by such a base is called the *product topology* or *Tychonov topology on* Ω. We notice that although the box topology is also a product topology on Ω, the Tychonov topology is most often referred to as the *product topology*. □

The product topology in Definition 5.7 was first introduced by the Russian mathematician Andrei Tychonov in some rudimentary form, in 1930. He generalized this concept to its present form in his later article, "Ein Fixpunktsatz" (see [Tych]). The same formalism of the product topology was also due to the Czech mathematician Eduard Čech (1893–1960) in 1937 who seemed to be unaware of Tychonov's papers.

Obviously, base $\mathcal{B}$ of (5.7 - 5.7a) for τ_P can be further reduced if each O_x is selected from a base $\mathcal{B}_x$ for τ_x. (See Problem 5.15.)

5.8 Remarks.

(i) Let $\mathcal{S}_x$ be a subbase for τ_x. One can show that the collection $\mathcal{S} = \{\pi_x^*(S_x) : S_x \in \mathcal{S}_x,\ x \in X\}$ of unit subbase cylinders is a subbase for τ_P, just as it took place in the case of finite products. (See Problem 5.16.)

(ii) We always prefer to deal with the smallest possible base or a subbase for τ_P, provided that we have the foreknowledge of τ_x's, such as their bases or subbases. For instance, as a rule of thumb, we can take $\{\pi_x^*(O_x) : O_x \in \tau_x\}$ as a subbase for τ_P, unless more is known about the nature of τ_x's and thus the subbase will be further reducible. □

5.9 Examples.

(i) Let $\{(Y_x, \tau_x) : x \in X\}$ be a collection of metrizable topological spaces and let $\Omega = \prod_{x \in X} Y_x$. According to Example 2.3 (i), the collection of all open balls $B_x(y_x, r),\ y_x \in Y_x,$ forms a base for $(Y_x, \tau_x(d_x))$. Now, the set of all simple cylinders of the type

$$\pi^*_{\{\alpha_1,\ldots,\alpha_k\}}\left(\prod_{i=1}^{k} B_{n_i}(y_i, r_i)\right)$$

$$= \pi^*_{\alpha_1}(B_{n_1}(y_1, r_1)) \cap \pi^*_{\alpha_2}(B_{n_2}(y_2, r_2)) \ldots \cap \pi^*_{\alpha_k}(B_{n_k}(y_k, r_k)),$$
$$\alpha_i \in X,\ y_i \in Y_{\alpha_i},\ k = 1, 2, \ldots, \tag{5.9}$$

forms a base for τ_P, whereas the collection of all unit cylinders of the type $\pi_x^*(B_{n_x}(y_x, r_x))$ must form a subbase for τ_P.

(ii) Let $\Omega = \mathbb{R}^{\mathbb{R}}$ be the collection of all real-valued functions on $\mathbb{R}$ regarded as the Cartesian product of $\mathbb{R}$'s, with each $\mathbb{R}$ endowed with the usual topology. We select an open neighborhood U_f of a point $f \in \Omega$. First of all, according to (5.9), a simple cylinder with base $(y_1 - \varepsilon_1, y_1 + \varepsilon_1) \times \ldots \times (y_k - \varepsilon_k, y_k + \varepsilon_k)$ has the form

$$\pi^*_{\{\alpha_1,\ldots,\alpha_k\}}\left(\prod_{i=1}^{k} (y_i - \varepsilon_i, y_i + \varepsilon_i)\right)$$

$$= \pi^*_{\alpha_1}[(y_1 - \varepsilon_1, y_1 + \varepsilon_1)] \cap \ldots \cap \pi^*_{\alpha_k}[(y_k - \varepsilon_k, y_k + \varepsilon_k)], \quad (5.9a)$$

where y_x is a point in $Y_x = \mathbb{R}$. In order that this cylinder becomes a neighborhood of f, we need to replace y_x with the corresponding traces $f(\alpha_x)$ of f in the factor spaces $Y_{\alpha_1}, \ldots, Y_{\alpha_k}$:

$$\pi^*_{\{\alpha_1,\ldots,\alpha_k\}}\left(\prod_{i=1}^{k}(f(\alpha_i) - \varepsilon_i, f(\alpha_i) + \varepsilon_i)\right). \quad (5.9b)$$

(See Figure 5.1.)

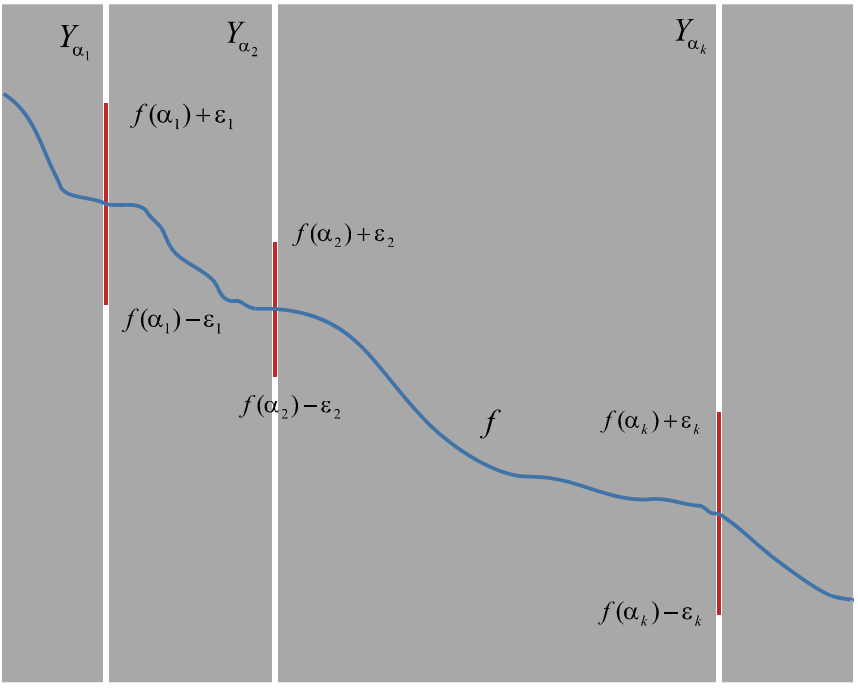


Figure 5.1

□

5.10 Remark. Let $\{g_x : x \in X\}$ be a family of functions $g_x : \Omega \to Y_x$, where each Y_x is endowed with a topology τ_x. Recall that $g^{**}_x(\tau_x)$, $\forall x \in X$, is a topology on Ω, and that each function g_x is

continuous relative to this topology. The union of all these topologies,

$$\mathcal{S} = \bigcup_{x\in X} g_x^{**}(\tau_x),$$

need not be a topology, for it does not necessarily preserve unions and intersections. But we can extend it to a topology, say $\tau(\mathcal{S})$, regarding $\mathcal{S}$ as a subbase. This topology is the weakest one relative to which all functions of the above family are continuous. Consequently, $\tau(\mathcal{S})$ is called the *weak topology generated by the family* $\{g_x\}$.

Now, taking $\prod_{x\in X} Y_x$ for Ω and π_x (the xth projection map) for g_x, by Problem 5.19, the Tychonov topology τ_P is the weakest topology relative to which all projection maps are continuous.

Furthermore, by Problem 5.16, $\bigcup_{x\in X} \pi_x^{**}(\mathcal{S}_x)$ is a subbase for τ_P. Other relevant subbases for τ_P that can be used are $\bigcup_{x\in X} \pi_x^{**}(\mathcal{B}_x)$ or $\bigcup_{x\in X} \pi_x^{**}(\tau_x)$. In a nutshell, τ_P turns out to be the weak topology generated by the projection maps. Consequently, this offers an alternative definition of the Tychonov topology on $\Omega = \prod_{x\in X} Y_x$. □

5.11 Example. Recall that a sequence $\{y_n\} \subseteq \Omega$ converges to a point $y \in \Omega$ if, for every neighborhood U_y, there is an $N(U_y)$-tail of $\{y_n\}$. Under the conditions of Example 5.9(ii), if Ω is the product space $\Omega = \mathbb{R}^{\mathbb{R}}$, a sequence of points $\{f_n\} \subseteq \Omega$ is convergent to a point $f \in \Omega$ if and only if $f_n(x) \to f(x)$ for each $x \in \mathbb{R}$. To see this we note that a base neighborhood U_f of f in (5.9b) is of the form,

$$U_f = \{g \in \Omega : |f(x_i) - g(x_i)| < \varepsilon_i\,,\ i = 1, \ldots, k\}, \qquad (5.11)$$

which is just another expression of (5.9b). Therefore, $f_n \to f$ if it is close to f on each finite set $\{x_1, \ldots, x_k\} \subseteq X = \mathbb{R}$, specifically on singletons $\{x\} \subseteq \mathbb{R}$. □

Example 5.11 motivates the following notion.

5.12 Definition. Let $\{(Y_x, \tau_x) : x \in X\}$, be an indexed family of topological space and let $\left(\prod_{x\in X} Y_x, \tau_P\right)$ be the Tychonov product topology. Recall (Remark 5.2 (ii), Chapter 1) that if $Y_x = Y$ and $\tau_x = \tau$, for each $x \in X$, then we denoted $\prod_{x\in X} Y_x$ by Y^X and called it the *set of functions from X to Y*. Now, in this particular case of the Tychonov product topology, in light of Example 5.11, (Y^X, τ_P) is called the *topology of pointwise convergence.* □

As a generalization of Example 5.11, the following proposition can help solidify our understanding of the topology of pointwise convergence.

5.13 Theorem. *Let $\{f_n\}$ be a sequence in Y^X. Then $f_n \to f \in Y^X$ (in the topology of pointwise convergence) if and only if $f_n(x) \to f(x)$, $\forall x \in X$ [in the topology (Y_x, τ_x)].*

Proof. Recall that $\pi_x : Y^X \to Y_x = Y$ is the x-projection map defined as $\pi_x(f) = f(x)$ [see Definition 5.4 (i), Chapter 1].

(i) First assume that $f_n \to f$ in (Y^X, τ_P). By Problem 5.19, π_x is continuous for every x. Thus, by Theorem 4.9, $\pi_x(f_n) \to \pi_x(f)$. This yields that $f_n(x) \to f(x)$ in (Y_x, τ_x).

(ii) Now assume that $f_n(x) \to f(x)$ in (Y_x, τ_x), $\forall x \in X$. We need to show that for any base neighborhood B_f of f, there is an $N(B_f)$-tail of $\{f_n\}$. Clearly, B_f is a simple cylinder of the form

$$B_f = \pi^*_{\{x_1,\ldots,x_k\}}\left(\prod_{i=1}^{k} O_{f(x_i)}\right),$$

where $O_{f(x_i)}$ contains $f(x_i)$.

By our assumption, $f_n(x_i) \to f(x_i)$, which implies the existence of an $N_i(O_{f(x_i)})$-tail, $i = 1, \ldots, k$. Let $N = \max\{N_1, \ldots, N_k\}$. (Note that this is exactly the place where we take advantage of the Tychonov product topology, for otherwise, in the case of the box topology, a base neighborhood of f could not be represented by a simple cylinder. The latter is an obstacle to finding a finite maximum of infinitely many N_i's implying that $\{f_n\}$ does not converge to f in

this box topology.) Then, for each x_i, $i = 1, \ldots, k$, we have the $N(O_{f(x_i)})$-tail of $\{f_n(x_i)\}$, which yields that

$$f_n \in \pi^*_{x_i}(\{f_n(x_i)\}) \subseteq \pi^*_{x_i}(O_{f(x_i)}),\ i = 1, \ldots, k.$$

Therefore, we have

$$f_n \in \bigcap_{i=1}^{k} \pi^*_{x_i}(f_n(x_i)) \subseteq \bigcap_{i=1}^{k} \pi^*_{x_i}(O_{f(x_i)}) = B_f, \text{ for all } n \geq N.$$

The latter yields the existence of the $N(B_f)$-tail of $\{f_n\}$, and therefore, $f_n \to f$ in (Y^X, τ_P). □

PROBLEMS

5.1 Let (Y_1, τ_1) and (Y_2, τ_2) be topological spaces, let (Ω, τ_P) be the product space, and let $F_i \in \tau_i^c, i = 1, 2$. Show that $F_1 \times F_2 \in \tau_P^c$.

5.2 Let $(Y_i, \tau_i), i = 1, \ldots, n,$ be topological spaces, (Ω, τ_P) be the product space, and let $F_i \in \tau_i^c, i = 1, \ldots, n$. Show that $\prod_{i=1}^{n} F_i \in \tau_P^c$.

5.3 Show that $Int(A \times B) = \mathring{A} \times \mathring{B}$.

5.4 Let $(Y_i, \tau_i), i = 1, \ldots, n,$ be topological spaces and (Ω, τ_P) be the product space. Show that $Int\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathring{A}_i$.

5.5 Let (Y_1, τ_1) and (Y_2, τ_2) be topological spaces, let (Ω, τ_P) be the product space, and let $A \subseteq Y_1$ and $B \subseteq Y_2$. Show that

$$\overline{A \times B} = \overline{A} \times \overline{B}.$$

5.6 Let $(Y_i, \tau_i), i = 1, \ldots, n,$ be topological spaces and (Ω, τ_P) be the product space. If $A_i \subseteq Y_i, i = 1, \ldots, n$, show that

$$\overline{\prod_{i=1}^{n} A_i} = \prod_{i=1}^{n} \overline{A}_i.$$

5.7 Let $(Y_i, \tau_i), i = 1, \ldots, n,$ be topological spaces and (Ω, τ_P) be the product space. Let $A_i \subseteq Y_i, i = 1, \ldots, n$. Is it true that

$$\partial\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \partial A_i?$$

5.8 Prove Proposition 5.1. *Let*

$$\widetilde{\mathcal{B}} = \mathcal{B}_1 \times \cdots \times \mathcal{B}_n : = \{B_1 \times \cdots \times B_n : B_i \in \mathcal{B}_i\}$$

where $\mathcal{B}_i$ is a base for τ_i, $i = 1, \ldots, n$. Then $\widetilde{\mathcal{B}}$ is also a base for τ_P.

5.9 Prove Proposition 5.2. *Let*

$$\mathcal{S} = \mathcal{S}_1 \times \cdots \times \mathcal{S}_n$$

$$= \{S_1 \times \cdots \times S_n : S_i \in \mathcal{S}_i,\ i = 1, \ldots, n\},$$

where $\mathcal{S}_i$ is a subbase for $\tau_i = \tau(\mathcal{S}_i)$, $i = 1, \ldots, n$. Then $\mathcal{S}$ is a subbase for τ_P.

5.10 Prove Proposition 5.3. *Let*

$$\mathcal{S}' = \{\pi_i^*(S_i) : S_i \in \mathcal{S}_i,\ i = 1, \ldots, n\},$$

where $\mathcal{S}_i$ is a subbase for τ_i. Then $\mathcal{S}'$ is a subbase for $\tau_P = \tau(\mathcal{S})$.

5.11 A map $f : (X, \tau) \to (Y, \tau')$ is said to be *open* if $f_{**}(\tau) \subseteq \tau'$. Show that in the product topology each projection map is open.

(*Hint:* Use the fact that, according to Problem 3.3, Chapter 1, maps preserve unions.)

5.12 Let $f:(\Omega,\tau)\to\left(\Omega=\prod_{i=1}^{n}Y_i,\tau_P\right)$ and $\pi_j:\Omega\to Y_j$. Show that the function f is continuous if and only if each $\pi_j\circ f$ is continuous. [*Hint*: Show that $f^*(S)\in\tau$, for every subbase element of τ_P, and then apply Theorem 4.8.]

5.13 Let (Y_i,τ_i) be a Hausdorff space, $i=1,\ldots,n$. Prove that $\left(\prod_{i=1}^{n}Y_i,\tau_P\right)$ is Hausdorff.

5.14 Let $(Y_x,\tau_x),x\in X$, be an indexed family of topological spaces and (Ω,τ_P) be the Tychonov product space. Let A be an arbitrary subset of X and $F_x\in\tau_x^c,x\in A$. Show that $\prod_{x\in A}F_x$ is closed in τ_P.

5.15 Let $\{(Y_x,\tau_x):x\in X\}$ be an indexed family of topological spaces and let

$$\widetilde{\mathcal{B}}=\left\{\pi_A^*\left(\prod_{x\in A}B_x\right):A\text{'s are finite subsets of }X,\ B_x\in\mathcal{B}_x\right\},$$

where $\mathcal{B}_x$ is a base for τ_x. Show that $\widetilde{\mathcal{B}}$ is also a base for the Tychonov topology τ_P on Ω.

5.16 Let $(Y_x,\tau_x),x\in X$, be an indexed family of topological spaces such that $\mathcal{S}_x$ is a subbase for $\tau_x,x\in X$. Show that the collection of all unit subbase cylinders

$$\mathcal{S}=\bigcup_{x\in X}\pi_x^{**}(\mathcal{S}_x)=\{\pi_x^*(S_x):S_x\in\mathcal{S}_x,x\in X\}$$

[in Remark 5.8(i) and Remark 5.10] is a subbase for the Tychonov topology.

5.17 Let $\{(Y_x, \tau_x) : x \in X\}$ be an indexed family of Hausdorff spaces. Prove that $\left(\prod_{x\in X} Y_x, \tau_P\right)$ is Hausdorff.

5.18 Let $\left(\Omega = \prod_{x\in X} Y_x, \tau_P\right)$ be the Tychonov topology and assume that each factor space is first countable. Is (Ω, τ_P) first countable if:

a) $|X| = \aleph_0$?
b) $|X| \geq \mathfrak{c}$?

5.19 Generalize Theorem 5.5 for the case of Tychonov's topology : *Let* $\Omega = \prod_{i\in X} Y_x$ *and* $\pi_x : \Omega \to Y_x$ *be the* x*th projection map. Then the Tychonov product topology* τ_P *on* Ω *is the weakest topology with respect to which each projection is continuous.*

5.20 Let $\left(\Omega = \prod_{x\in X} Y_x, \tau_P\right)$ be the Tychonov topology and assume that each factor space is second countable. Is (Ω, τ_P) second countable if X is countable?

5.21 Let $\left(\Omega = \prod_{x\in X} Y_x, \tau_P\right)$ be the Tychonov topology of (Y_x, τ_x) factor spaces. If $A_x \subseteq Y_x, x \in X$, show that

$$\overline{\prod_{x\in X} A_x} = \prod_{x\in X} \overline{A}_x.$$

5.22 Let $\left(\Omega = \prod_{x\in X} Y_x, \tau_P\right)$ be the Tychonov topology and assume that each factor space is separable. Is (Ω, τ_P) separable if X is countable?

NEW TERMS:

$\mathcal{B} = \underset{i=1}{\overset{n}{\times}} \tau_i$, base for product topology, 213
product topology of finitely many factor topologies, 213
base (simple) rectangle in product topology, 213
simple (base) rectangle in product topology, 213
$\underset{i=1}{\overset{n}{\times}} \mathcal{S}_i$, subbase for τ_P, 213
subbase rectangle, 214
continuity of projection maps, 215, 224
product topology of arbitrarily many factor topologies, 215
box topology, 215
box rectangle, 215
Tietze, Heinrich, 215
$\left(\prod_{x \in X} Y_x, \tau_P \right)$, Tychonov product topology, 216
Tychonov, Andrei, 216
Čech, Eduard, 216
open neighborhood of f, 218
weak topology (generated by a family of functions), 219
(Y^X, τ_P), topology of pointwise convergence, 220
pointwise convergence, criterion of, 220
open map, 222

6. NOTES ON SUBSPACES AND COMPACTNESS

It has been mentioned that subspaces of topological spaces (that is, relative topologies) inherit certain features of the original spaces. In this section we consider this notion more systematically. We are concerned with such topological properties as separability, countability, and compactness, and their impact on subspaces.

6.1 Definition. A property of a space is referred to as *hereditary* if every subspace has this property. A property is said to be *weakly hereditary* if it is inherited by a subspace whose carrier is closed in the original space. A property is *vaguely hereditary* if it is inherited by a subspace whose carrier is open in the original space. (The last notion is restricted to the use in this textbook.) □

6.2 Example. Second countability is hereditary. (See Problem 3.7.) □

6.3 Remark. In Section 1 we denoted by $\bar{A}$ the closure of some subset A of a topological space (X, τ), understanding that this is the closure relative to the topology τ. As mentioned in Definition 1.8(i), in the case of subspaces we may need to deal with closures of subsets with respect to any relative topology, say (Y, τ_Y). To make a certain distinction clear we then write $Cl_Y A$. However, we will still use $\bar{A}$ whenever we have in mind the closure relative to the original space (X, τ). □

6.4 Example. The property of density of a set is not hereditary and not weakly hereditary, that is, if D is dense in (X, τ), its trace in a subspace (Y, τ_Y) need not be dense. Let $(X, \tau) = (\mathbb{R}, \tau_e)$ and $Y = \mathbb{R}_+ \cup \{-\sqrt{2}\}$. Then, obviously the set $\mathbb{Q}_+ = \mathbb{Q} \cap Y$ is not dense in (Y,τ_Y). It is easily seen that $\{-\sqrt{2}\}$ is an open neighborhood of the point $-\sqrt{2}$ that does not meet $\mathbb{Q}_+$. Thus $Cl_Y \mathbb{Q}_+ \neq Y$. Because Y is closed in $(\mathbb{R}, \tau_e)$, the density property is not weakly hereditary either. □

6.5 Theorem. *Separability is vaguely hereditary, but is neither hereditary nor weakly hereditary.*

Proof.

(i) Let (X, τ) be separable and let (Y, τ_Y) be a subspace of (X, τ) such that $Y \in \tau$. We show that (Y, τ_Y) is separable. Let D be a countable dense set in (X, τ). We need to prove that $Cl_Y(D \cap Y) = Y$; specifically, we need to show that $Y \subseteq Cl_Y(D \cap Y)$, for the inverse inclusion holds trivially. Let y be any point of Y and let U'_y be any open neighborhood of y in τ_Y. Because Y is open in X, U'_y is also a neighborhood of the point y in τ. (It is easy to show the following. Since U'_y is a neighborhood of y in τ_Y, there is $O'_y \subseteq U'_y$ which is τ_Y-open. But $O'_y = O_y \cap Y$, where $O_y \in \tau$. Because Y is τ-open, it follows that O'_y is also τ-open and, clearly, $U'_y \in \mathcal{U}_y$ in τ.)

Therefore, U'_y meets D and, consequently, U'_y meets $D \cap Y$ (as a subset of Y). Observe that if Y is not open in X, U'_y need not be a neighborhood of y in Y. [For instance, let $Y = (0, 2]$ and $U'_2 = (1, 2]$. Clearly, U'_2 is not a neighborhood of 2 in $(\mathbb{R}, \tau_e)$, but it is a neighborhood of 2 in (Y, τ_Y).]

(ii) As a counterexample of separability as a hereditary property, we consider the topology (X, τ) known as the *Moore plane*. Let $X = \mathbb{R} \times [0, \infty)$ (the upper semiplane with the horizontal axis). The topology on X will be described by the following base sets. At each point $(x, y) \in \mathbb{R} \times (0, \infty)$, a neighborhood base is the collection of all open balls $\{B((x, y), r) : r < y\}$, where $B(z, r) = \{z' \in X : d_e(z, z') < r\}$. At each point $(x, 0)$, a neighborhood base consists of all open balls touching the horizontal axis at $(x, 0)$, and the point $(x, 0)$ is attached to these balls.

Take the union of all neighborhood bases and construct a base for the Moore plane topology in light of Theorem 2.2. This topological space is separable with the dense countable subset $D = \mathbb{Q}^2 \cap X$. Indeed, let $(x, y) \in \mathbb{R} \times (0, \infty)$. Then any neighborhood of (x, y) contains points with rational coordinates (a property inherited from the Euclidean space). As for the points $(x, 0)$, any open ball bordering $(x, 0)$ also contains points with rational coordinates.

Now, for a subspace of the Moore plane, consider one with the horizontal axis as the carrier Y. Clearly, all singletons are traces of

base neighborhoods at $(x, 0)$ in Y yielding the discrete topology as the relative topology on Y. According to Problem 6.2, any discrete topology with an uncountable carrier is not separable. Observe that Y^c is obviously open in X. Hence the separability is not weakly hereditary. □

6.6 Definition. A subset A of a topological space (X, τ) is said to be *compact* (or *Lindelöf compact* or just *Lindelöf*) if every open cover of A contains a finite (at most countable) subcover. We also say that A is *finitely* (*countably*) *reducible*. Specifically, if X is compact (Lindelöf), (X, τ) is called a *compact topological space* (*Lindelöf space*). □

6.7 Example. Compactness in metrizable topological spaces obviously coincides with that for the corresponding metric spaces. In this case, we may use the tools and criteria of compactness for metric spaces. For instance, the interval $[a, b]$ for $a, b \in \mathbb{R}^n$ is compact in the sense of the Euclidean metric; thus, it is compact in $(\mathbb{R}^n, \tau_e)$, whereas (a, b) is not compact in $(\mathbb{R}^n, \tau_e)$, because it is not closed. □

Theorem 6.8 below is almost identical to Theorem 6.10 of Chapter 2. However, for consistency, we formulate it and prove it again.

6.8 Theorem. *Let $f : (X, \tau) \to (Y, \tau')$ be a continuous function and let K be a compact subset of X. Then the image $f_*(K)$ is compact. In short, the image of a compact set under a continuous function is compact.*

Proof. Take any open cover $\{O'_i : i \in I\}$ of $f_*(K)$ to have $f_*(K) \subseteq \bigcup_{i\in I} O'_i$. Then,

$$K \subseteq f^*(f_*(K)) \subseteq \bigcup_{i\in I} f^*(O'_i).$$

Because f is continuous, $f^*(O'_i)$ is open, and because K is compact, there is a finite subcover of sets $f^*(O'_i)$, without loss of generality indexed by $1, \ldots, n$, that is,

$$K \subseteq \bigcup_{k=1}^{n} f^*(O'_k).$$

Because by Problem 3.6(a, b), Chapter 1, maps preserve inclusions and unions,

$$f_*(K) \subseteq \bigcup_{k=1}^{n} f_*(f^*(O'_k)) = \bigcup_{k=1}^{n} O'_k. \qquad \square$$

6.9 Theorem. *Compactness is weakly hereditary* (*that is*, *a closed subset of a compact topological space is compact*).

Proof. Let (X, τ) be compact and let $B \subseteq X$ be closed. Let $\{O_i : i \in I\}$ be an open cover of B. Because B^c is open, $\{B^c, O_i : i \in I\}$ is an open cover of X. Because X is compact, there exists a finite open subcover of X, say $\{B^c, O_1, \ldots, O_n\}$, which is also an open subcover of B. Hence, B is compact. $\square$

Hausdorff topological spaces have an important property with respect to compactness.

6.10 Theorem. *Every compact subset of a Hausdorff space is closed.*

Proof. Let A be a compact subset of the Hausdorff space (X, τ). We show that A^c is open. Take $x \in A^c$. The family of neighborhoods of all points $y \in A$ covers A. We extract a particular subfamily of these neighborhoods. Because (X, τ) is Hausdorff, for each $y \in A$, there is a neighborhood $U_x(y)$ of x and a neighborhood $V_y(x)$ of y such that $U_x(y) \cap V_y(x) = \emptyset$. Without loss of generality we may assume that the family $\{V_y(x) : y \in A\}$ is an open cover of A. Because A is compact, there exists an open subcover $\{V_{y_k}(x) : k = 1, \ldots, n\}$ of A. Obviously, $V_{y_k}(x) \cap U_x(y_k) = \emptyset$. Select $\{O_x(y_k) \subseteq U_x(y_k),\ k = 1, \ldots, n\}$, whose intersection (denoted by O_x), because being finite, is open and nonempty. Therefore, O_x is an open neighborhood of $x \in A^c$ with $O_x \cap A = \emptyset$, which means that x is an interior point of A^c. Thus, A^c is open, or equivalently, A is closed. $\square$

6.11 Remark. In Theorem 6.3, Chapter 2, we stated and proved (Problem 6.8) the equivalence of the conditions: (i) $A \subseteq (X, d)$ is compact; (ii) every infinite subset of A has an accumulation point in A (*Bolzano - Weierstrass compactness*); and (iii) every sequence in

A has a convergent subsequence (*sequential compactness*). This equivalence does not hold for topological spaces, where (i) and (iii) are in general distinct properties, and compactness just implies Bolzano - Weierstrass compactness, as the reader proves (see Problem 6.6). □

Recall that second countable spaces are first countable and separable (see Problem 3.6). In addition, they are Lindelöf spaces, as the following theorem asserts.

6.12 Theorem. *Any second countable topological space is Lindelf*□.

Proof. Let $\mathcal{B}$ be a countable base of a topological space (X, τ) and let $O = \{O_i : i \in I\}$ be an arbitrary open cover of X. Let $x \in X$. Then x belongs to some $O_x \in O$. Because $\mathcal{B}$ is a base for τ, by Theorem 2.2, there is a neighborhood base $\mathcal{B}_x \subseteq \mathcal{B}$. Then there is a base neighborhood $B_x \in \mathcal{B}_x$ such that $B_x \subseteq O_x$. The collection of all distinct B_x's for all $x \in X$ is at most countable. On the other hand, this collection obviously covers X. Consequently, the collection of all open supersets $\{O_x\}$ associated with each B_x is also countable and it covers X. Thus (X, τ) is indeed Lindelöf. □

The result below is in the spirit of Theorem 4.8, where, for continuity of a function f from (X, τ) to $(Y, \tau(\mathcal{S}))$, it was sufficient to verify that $f^*(\mathcal{S}) \subseteq \tau$. Here we claim that, if $\mathcal{S}$ is a subbase for τ, then X is compact whenever every cover of X by elements from $\mathcal{S}$ can be reduced to a finite subcover.

6.13 Theorem (**Alexander Subbase Theorem**). *Let (X, τ) be a topological space and let σ be a subbase for τ. If every open cover of X by elements of σ is finitely reducible, that is, if every open cover can be reduced to a finite subcover, then X is compact.*

Proof. We prove the equivalent statement. *If X is not compact, then there exists an open cover by elements of σ that is not finitely reducible.*

Assume that X is not compact. We prove this assertion in four steps, which we outline as follows.

(i) Let $\mathbb{O}$ be the collection of all open covers of X that cannot be reduced to finite subcovers. We show that $\mathbb{O}$ has a maximal element; call it $\mathcal{M}$.

(ii) We show that for every $x \in X$, there is an open set $M_x \in \mathcal{M}$ and a finite tuple of open sets $\{S_1(x), \ldots, S_n(x)\} \subseteq \sigma$ such that

$$x \in \bigcap_{k=1}^{n} S_k(x) \subseteq M_x.$$

(iii) We show that at least one of the sets $\{S_1(x), \ldots, S_n(x)\}$, denote it $S(x)$, belongs to $\mathcal{M}$.

(iv) We recognize that for each $x \in X$, $S(x) \in \sigma$ and $S(x) \in \mathcal{M}$. In particular, the latter implies that $\{S(x) : x \in X\}$ is an open cover of X, which is not finitely reducible. On the other hand, because we will have $\{S(x) : x \in X\} \subseteq \sigma$, the proof be complete.

We are be concerned with each of the above steps in detail.

Step (i): Because X is not compact, $\mathbb{O}$ is not empty. Introduce on $\mathbb{O}$ the partial order relation in terms of the inclusion. (In other words, two open covers, $\mathcal{C}_1$ and $\mathcal{C}_2$ of X from $\mathbb{O}$ are related as $\mathcal{C}_1 \subseteq \mathcal{C}_2$ if and only if $\mathcal{C}_1$ is a subcover of the cover $\mathcal{C}_2$.) Let $\mathbb{C} \subseteq \mathbb{O}$ be any chain, and let $\mathcal{U}$ be the union of all elements of $\mathbb{C}$. Clearly, $\mathcal{U}$ is a cover of X that cannot be reduced to a finite subcover. Thus, $\mathcal{U}$ must belong to $\mathbb{O}$, and $\mathcal{U}$ is an upper bound of $\mathbb{C}$. By Zorn's Lemma 4.13, Chapter 1, there is at least one maximal element in $\mathbb{O}$; denote it by $\mathcal{M}$.

Step (ii): Let $x \in X$ and let $M_x \in \mathcal{M}$ such that x is an interior point of M_x (which exists, for $\mathcal{M} \in \mathbb{O}$ is an open cover). On the other hand, by Theorem 2.7, the collection $\mathcal{B}$ of all finite intersections of elements of the subbase σ is a base $\mathcal{B}$ for τ. Thus, as an open set, M_x is a union of base elements, each one of which is a finite intersection of elements of σ; that is, x belongs to one of the base elements B_x, represented by a finite intersection $\bigcap_{k=1}^{n} S_k(x)$ of

subbase elements. In other words, there is a tuple $\{S_1(x), \ldots, S_n(x)\} \subseteq \sigma$ such that $x \in \bigcap_{k=1}^{n} S_k(x) \subseteq M_x$.

Step (iii): Assume that for the given set M_x there is no element of the tuple $\{S_1(x), \ldots, S_n(x)\}$ which is an element of $\mathcal{M}$. In this case, for each $S_k(x) \in \{S_1(x), \ldots, S_n(x)\}$, there is a finite subcollection $\{M_1(k), \ldots, M_{j_k}(k)\}$ from $\mathcal{M}$ that supplements $S_k(x)$ to a cover of X. If no such finite subcollection were to exist, then $\{S_k(x)\} \cup \mathcal{M}$ would be an element of $\mathbb{O}$, that is, $\{S_k(x)\} \cup \mathcal{M}$ would not be finitely reducible, and $\mathcal{M}$ would not be a maximal element of $\mathbb{O}$. Hence, $\{M_1(k), \ldots, M_{j_k}(k), S_k(x)\}$, $k = 1, \ldots, n$, is a finite open cover of X. By Problem 6.9,

$$\{M_1(k), \ldots, M_{j_k}(k) : k = 1, \ldots, n, \text{ and } \bigcap_{k=1}^{n} S_k(x)\}$$

is also a finite open cover of X. This implies that

$$\{M_1(k), \ldots, M_{j_k}(k) : k = 1, \ldots, n \text{ , and } M_x\}$$

is a finite cover of X, and is also a finite subcover of $\mathcal{M}$, which contradicts the property that $\mathcal{M}$ is an element of $\mathbb{O}$. Thus, our assumption about $\{S_1(x), \ldots, S_n(x)\}$ was wrong and there is at least one set $S(x) \in \{S_1(x), \ldots, S_n(x)\}$ that belongs to $\mathcal{M}$.

Step (iv): It follows that, for each $x \in X$, there is an element $S(x)$ common to σ and $\mathcal{M}$, and thus the collection $\{S(x) : x \in X\}$ is an open cover of X that obviously cannot be reduced to a finite subcover. (Observe that the assumption that X is not compact implies that X cannot be finite.) □

The Alexander subbase theorem leads to the following meaningful result by Tychonov.

6.14 Theorem (**Tychonov**). *A nonempty Tychonov product* $\left(\Omega = \prod_{x \in X} Y_x, \tau_P\right)$ *is compact if and only if each factor space* (Y_x, τ_x) *is compact.*

Proof.

(i) If the Tychonov product is compact, then compactness of factor spaces follows from the continuity of projection maps (see Problem 5.19 and Theorem 6.8).

(ii) Let (Ω, τ_P) be the Tychonov product of compact spaces (Y_x, τ_x), $x \in X$. Take for a subbase for τ_P the collection of all simple cylinders $\mathcal{S} = \{\pi_x^*(O_x) : O_x \in \tau_x,\ x \in X\}$ (see Proposition 5.3). If Ω is not compact, by the Alexander subbase theorem, no subcollection of $\mathcal{S}$ covering Ω can be reduced to a finite subcover. Specifically, $\mathcal{S}$ itself cannot be reduced to a finite subcover. Let $\{O_x\}$ be an arbitrary open cover of Y_x. Then it can be reduced to a finite subcover, $\{O_x(1), \ldots, O_x(k_x)\}$ of Y_x . Thus, $\{\pi_x^*(O(1)), \ldots, \pi_x^*(O_x(k_x))\}$ is a finite open cover of Ω that is a finite subcover of the largest cover, $\mathcal{S}$. This contradicts the hypothesis that Ω is not compact. □

PROBLEMS

6.1 Is the property of a set to be dense vaguely hereditary? (Consider Example 6.4.)

6.2 Let $(X, \mathcal{P}(X))$ be a discrete topological space with an uncountable carrier. Show that the space is not separable.

6.3 Show that the topological space in Problem 6.2 is not second countable.

6.4 Show that first countability is hereditary.

6.5 Prove that the Moore plane is not second countable.

6.6 Prove the statement. *Every compact topological space is Bolzano - Weierstrass compact*.

6.7 Let τ be the cofinite topology on an arbitrary nonempty set X. Show that (X, τ) is compact.

6.8 Show that any bijective continuous map $f\colon (X, \tau) \to (Y, \tau')$, where (X, τ) is compact and (Y, τ') is Hausdorff, is a homeomorphism. (*Hint*: Make use of Proposition 4.3.)

6.9 Let $\{M_1(k), \ldots, M_{j_k}(k), S_k\}$ be a cover of a set X for each $k = 1, \ldots, n$. Show that $\{M_1(k), \ldots, M_{j_k}\colon\ k = 1, \ldots, n\,,$ and $\bigcap_{k=1}^{n} S_k\}$ is also a cover of X.

6.10 Let (X, τ) be a Hausdorff topological space. Show that an arbitrary intersection of compact sets in X is compact.

6.11 Let O be an open set in $\mathbb{R}^n$. Show that there is a monotone increasing sequence $\{O_k\}\uparrow$ of open bounded subsets of O such that $\{O_k\}\uparrow O$.

6.12 Is Hausdorffness hereditary in a strong or any weak sense?

6.13 Let (X, τ) be a Hausdorff space and C and K be disjoint compact sets. Show that there are disjoint open supersets, U and V, of C and K, respectively.

6.14 A topological space (X, τ) is called *countably compact* if any open countable cover of X has a finite subcover. Prove that the following are equivalent.

(i) (X, τ) is countably compact.

(ii) Each countable family of closed sets in (X, τ) with the finite intersection property (that is, the intersection of any finite subfamily is nonempty) has a nonempty intersection.

(iii) Every countably infinite subset A of X has a point x with the property that each neighborhood of x contains infinitely many points of A.

(iv) Every sequence in X has a closure point.

6.15 Let $\mathcal{K}$ be the family of all compact subsets of a topological space (X, τ). Under the formation of which set operations is $\mathcal{K}$ closed? Give counterexamples of those set operations, which are not contained in $\mathcal{K}$.

NEW TERMS:

hereditary property in topological spaces, 226
weakly hereditary property in topological spaces, 226
vaguely hereditary property in topological spaces, 226
Moore plane, 227
compact set in topological spaces, 228
Lindelöf compact set in topological spaces, 228
compact topological space, 228
Lindelöf topological space, 228
compact set under a continuous map in a topological space, 228
compact sets in Hausdorff spaces, 229
Bolzano - Weierstrass compactness in topological spaces, 229-230
sequential compactness in topological spaces, 230
Alexander subbase theorem, 230
Tychonov's theorem, 232
countably compact topological space, 234

Chapter 4

Measurable Spaces and Measurable Functions

In the previous chapter we studied general topological spaces. A topology was defined as a collection of sets (on a carrier) that is closed with respect to the formation of arbitrary unions and finite intersections. In the present chapter, we introduce various classes of sets similar to topological spaces but serving other purposes. One of them prepares the reader to another part of analysis, integration. Beyond the familiar integration we experienced in calculus, we need to measure much more general sets than those used for the Riemann integral. For instance, we consider abstract sets that are encountered in the theory of probability. In addition, we largely extend the existing class of integrable functions.

If we try to measure the length (or area) of all sets, set theory drives us into certain contradictions or paradoxes. Therefore, we have to restrict our attention to measuring a subclass of sets. It stands to reason that we would want the collection of "measurable" sets to be closed under certain operations such as union, complementation, and intersection. Thus we seek a collection of sets satisfying certain algebraic properties under the binary operations of union, intersection, and set-theoretic difference. This leads to the concept of a sigma-algebra.

As with topological spaces, where base (or sometimes also subbase) sets were most convenient to study, in measure theory it is also useful to start with more primitive collections of sets called *generators* of sigma-algebras. For instance, if we need to measure a

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0_4, © Springer Science+Business Media, LLC 2013

flat closed figure, one of the reasonable ways to render it is to approximate the figure by a number of specified (disjoint) rectangles whose measures we already know. Such a natural way of measuring more complex sets by "base" sets gives rise to the extension of measure from the collection of "abstract rectangles" to the set of all figures formed from rectangles under countable set operations. This way of extension was formalized by the German mathematician Constantin Carathéodory in 1918.

The present chapter is just preparation for the forthcoming two chapters, where we are concerned with various classes of sets on which measure will be defined and then extended. We will be introduce to some generators of these classes, in particular, topologies. The latter are of special interest.

1. SYSTEMS OF SETS

1.1 Definition. Let Ω be an arbitrary set and let Σ be some collection of subsets of Ω.

(i) Σ is called a *σ-algebra* (pronounced "sigma-algebra") or *σ-field* (sigma-field) if:

(a) $\Omega \in \Sigma$.

(b) $A \in \Sigma \;\Rightarrow\; A^c \in \Sigma$.

(c) For any sequence $\{A_n\}$ of sets of Σ, $\bigcup_{n=1}^{\infty} A_n \in \Sigma$.

(ii) Σ is called an *algebra* (or field), also denoted by $\mathcal{A}$ (that is, $\Sigma = \mathcal{A}$) if

(a) $\Omega \in \mathcal{A}$.

(b) $A \in \Sigma \;\Rightarrow\; A^c \in \mathcal{A}$.

(c) $\{A_1, \ldots, A_k\} \subseteq \Sigma \;\Rightarrow\; \bigcup_{n=1}^{k} A_n \in \mathcal{A}$.

(iii) Σ is called a *Dynkin system*, denoted by $\mathcal{D}$ ($\Sigma = \mathcal{D}$), if

(a) $\Omega \in \mathcal{D}$.

(b) $A \in \mathcal{D} \Rightarrow A^c \in \mathcal{D}$.

(c) For any sequence $\{A_n\}$ of pairwise disjoint sets of $\mathcal{D}$,

$$\sum_{n=1}^{\infty} A_n \in \mathcal{D}.$$

(iv) Σ is called a *ring*, denoted by $\mathcal{R}$ ($\Sigma = \mathcal{R}$), if

(a) $\emptyset \in \mathcal{R}$.

(b) $A, B \in \mathcal{R} \Rightarrow A \backslash B \in \mathcal{R}$.

(c) $A, B \in \mathcal{R} \Rightarrow A \cup B \in \mathcal{R}$.

(v) Σ is called a *semi-ring*, denoted by $\mathcal{S}$ ($\Sigma = \mathcal{S}$), if

(a) $\emptyset \in \mathcal{S}$.

(b) $A, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S}$.

(c) for $A, B \in \mathcal{S}$, there is a finite tuple $C_1, \ldots, C_k$ of pairwise disjoint sets from $\mathcal{S}$ such that $A \backslash B$ can be represented as the union $\sum_{n=1}^{k} C_n$.

(vi) Σ is called a *monotone system*, denoted by $\mathcal{M}$ ($\Sigma = \mathcal{M}$), if:

(a) for every $\{A_n\}\uparrow$ (that is, monotone nondecreasing) sequence of sets of $\mathcal{M}$,

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}.$$

(b) for every $\{A_n\}\downarrow$ (that is, monotone nonincreasing) sequence of sets of $\mathcal{M}$,

$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{M}.$$

(vii) $\cap$-*stable* if $A, B \in \Sigma \Rightarrow A \cap B \in \Sigma$.

$(viii)$ A pair (Ω, Σ), where Σ is a σ-algebra in Ω, is called a *measurable space*, and elements of Σ are called *measurable sets*. (Compare these with topological space and open sets, respectively.) $\square$

1.2 Examples.

(i) Let Ω be an arbitrary set. Then $\{\Omega, \emptyset\}$ is the smallest algebra (σ-algebra).

(ii) $\mathcal{P}(\Omega)$, the power set, is always a σ-algebra. It is the largest σ-algebra in Ω.

(iii) The smallest σ-algebra containing a set A is obviously $\{\Omega, \emptyset, A, A^c\}$.

(iv) Let $\Omega = \mathbb{R}^n$ and let $\mathcal{S}$ be the system of all n-dimensional half-open intervals (or *rectangles*) of type $(a, b]$, for $a, b \in \mathbb{R}^n$. The intersection of two intervals is either $\emptyset$ or again an interval. The difference of two intervals need not be an interval, but it can be represented as a finite union of pairwise disjoint intervals (see Figure 1.1). Hence, $\mathcal{S}$ is a semi-ring.

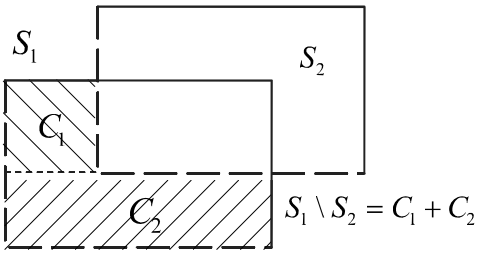


Figure 1.1

(v) Every σ-algebra Σ has three properties directly following from Definition 1.1(i):

a) $\emptyset \in \Sigma$ (because $\emptyset = \Omega^c$).

b) For every sequence $\{A_n\} \subseteq \Sigma$, $\bigcap_{n=1}^{\infty} A_n \in \Sigma$. This holds because

$$\{A_n\} \subseteq \Sigma \;\Rightarrow\; \{A_n^c\} \subseteq \Sigma \;\Rightarrow$$

$$\bigcup_{n=1}^{\infty} A_n^c \in \Sigma \;\Rightarrow\; \Big(\bigcup_{n=1}^{\infty} A_n^c\Big)^c \in \Sigma \;\Rightarrow\; \bigcap_{n=1}^{\infty} A_n \in \Sigma.$$

Observe that by applying De Morgan's law we can similarly show that this property and property (c) in Definition $1.1(i)$ are equivalent, that is, in $1.1(i, c)$ a countable union can be replaced with a countable intersection.

c) Finally we have

$$A, B \in \Sigma \;\Rightarrow\; A\backslash B \in \Sigma \text{ (due to } A\backslash B = A \cap B^c).$$

One can say that any σ-algebra is closed with respect to the formation of all countable set operations.

(vi) Every algebra $\mathcal{A}$ has the same property as σ-algebras in (v) except it is closed under finite intersections. Hence, any algebra is closed relative to the formation of all finite set operations.

(vii) Let Ω be an arbitrary set and let Σ be the system of all subsets A of Σ such that either A or A^c is at most countable. Then Σ is a σ-algebra. Indeed, $\emptyset$ is at most countable. Thus, $\Omega \in \Sigma$. If $A \in \Sigma$ then obviously $A^c \in \Sigma$. Now let $\{A_n\} \subseteq \Sigma$. Then either there is at least one countable set A_i, or else all A_n are not countable but their complements are countable. In the first case, $\bigcap\limits_{n=1}^{\infty} A_n$ is clearly countable. In the second case, the set $\bigcup\limits_{n=1}^{\infty} A_n^c$ is at most countable and, therefore, it belongs to Σ, along with its complement. The latter, by De Morgan's law, is $\bigcap\limits_{n=1}^{\infty} A_n$. Consequently, Σ is a σ-algebra.

$(viii)$ Let $f\colon \Omega \to \Omega'$ be a map and Σ' a σ-algebra in Ω'. Then $\Sigma = f^{**}(\Sigma')$ defined as the system of all sets $\{f^*(A')\colon A' \in \Sigma'\}$ is a σ-algebra. This property is due to the fact that the inverse of a map preserves all set operations. For instance, if $A \in \Sigma$, then it follows from the definition of Σ' that A is the inverse image of some $A' \in \Sigma'$. Thus $(A')^c \in \Sigma'$ yielding $A^c = (f^*(A'))^c = f^*((A')^c) \in \Sigma$. The proof that the union of a sequence from Σ belongs to Σ is similar (show it).

(ix) A monotone system need not be an algebra. For instance, let $\mathfrak{C}$ be the set of all convex subsets of $\mathbb{R}^2$. Then it is easily seen that $\mathfrak{C}$ is a monotone system. However, the union of two convex sets $A \cup B$

need not be convex. The difference of two convex sets need not be convex either. On the other hand, an algebra need not be a monotone system, for it is not closed under countable set operations.

(x) The collection of all finite subsets of an infinite set Ω is a ring in Ω. □

PROBLEMS

1.1 Let Σ be a σ-algebra in Ω and let Ω' be a subset of Ω. Define Σ' as $\Sigma \cap \Omega' = \{A \cap \Omega' : A \in \Sigma\}$. Then Σ' (the *trace of* Σ *in* Ω') is a σ-algebra in Ω', called a *trace* σ*-algebra* . Prove it.

1.2 Show that any ring is closed under the formation of finite intersections.

1.3 Let $\mathcal{R}$ be a collection of subsets of Ω with the properties:

$$\emptyset \in \mathcal{R} \text{ and } A, B \in \mathcal{R} \Rightarrow A \cap B \in \mathcal{R} \text{ and } A \cup B \in \mathcal{R}.$$

Is $\mathcal{R}$ a ring?

1.4 Let $\mathcal{A}$ be an algebra in Ω. Prove that $\mathcal{A}$ is a σ-algebra in Ω if and only if $\mathcal{A}$ is also a monotone system.

1.5 The flowchart below reflects the relations between some systems of sets

$$\begin{array}{l} \mathcal{D} \\ \Uparrow \\ \Sigma \Rightarrow \mathcal{A} \Rightarrow \mathcal{R} \Rightarrow \mathcal{S} \Rightarrow \cap\text{-stable} \\ \Downarrow \\ \mathcal{M} \end{array}$$

Demonstrate that each relationship holds.

1.6 Give an infinite collection of subsets of $\mathbb{R}$ that contains Ø and $\mathbb{R}$ and which is closed under countable unions and intersections but is not a σ-algebra.

1.7 Let Ω be a finite set with $|\Omega| = 2n$. Let $\mathcal{D}$ be the system of all subsets D of Ω such that $|D| = 2q,\ q = 0, 1, \ldots, n$. Show that $\mathcal{D}$ is a Dynkin system.

1.8 Give a Dynkin system that is not a σ-algebra.

1.9 Show that if $\mathcal{D}$ is a Dynkin system then $D, E \in \mathcal{D}$ and $E \subseteq D \Rightarrow D \backslash E \in \mathcal{D}$.

1.10 Prove the statement: *A Dynkin system $\mathcal{D}$ is a σ-algebra if and only if $\mathcal{D}$ is $\cap$-stable.*

1.11 Show that the inverse image of a ring $\mathcal{R}$ under a map $f : \Omega_1 \to \Omega$ is a ring in Ω_1.

1.12 Show the equivalence of two definitions of a semi-ring if property c) in Definition 1.1(v) can be replaced with

c') Let $A, A_1, \ldots, A_n \in \mathcal{S}$. Then there is a finite tuple $C_1, \ldots, C_k$ of disjoint sets from $\mathcal{S}$ such that

$$A \backslash \bigcup_{s=1}^{n} A_s = \sum_{j=1}^{k} C_j.$$

NEW TERMS:

Carathéodory, Constantin, 240
σ-algebra (σ-field), 240
σ-field, 240
algebra (field), 240
field, 240
Dynkin system, 240
ring, 241
semi-ring, 241
monotone system, 241
$\cap$-stable (intersection-stable) system, 241
(Ω, Σ), measurable space, 241
measurable sets, 241
smallest σ-algebra, 242
σ-algebra containing (generated by) a set, 242
half-open interval (rectangle), 242
systems of sets, diagram of, 244

2. SYSTEMS' GENERATORS

This section is concerned with generators of σ-algebras which is somewhat analogous to the way how subbases and bases generate topologies. However, in the case of σ-algebras, there will be more varieties of generators including topologies themselves. We begin with the following theorem.

2.1 Theorem. *The intersection of arbitrarily many σ-algebras* (*algebras*, *monotone systems*, *rings*) *in Ω is a σ-algebra* (*an algebra*, *monotone system*, *ring*). □

(See Problem 2.1 .)

2.2 Remark. Let $\mathcal{G}$ be an arbitrary collection of subsets of Ω. There is obviously a σ-algebra containing $\mathcal{G}$, for instance, the power set $\mathcal{P}(\Omega)$. If we collect all σ-algebras that contain $\mathcal{G}$ and find their intersection, it must contain $\mathcal{G}$ and, due to Theorem 2.1, it is a σ-algebra too. This σ-algebra is clearly the smallest one containing $\mathcal{G}$. It is called the *σ-algebra generated by $\mathcal{G}$* and it is denoted by $\Sigma(\mathcal{G})$. The system of sets $\mathcal{G}$ is called the *generator* of $\Sigma(\mathcal{G})$. It is worthwhile to recollect the analogue of a subbase or base and their role as generators of the smallest topology. Although, as we learned, the classes of generators of topologies are somewhat limited, their counterparts for σ-algebras are more functional and form a significantly richer reserve filled with such prominent systems as semi-rings, rings, Dynkin systems, and monotone systems. Among them, rings and semi-rings are often used as generators of σ-algebras (throughout this book, especially, in Chapter 5) in Carathéodory construction of measures. Another frequently used generator is a topology itself that we see in action regarding a characterization of regular and Radon measures [Dsh2] and in calculus of Lebesgue - Stieltjes measures. The smallest σ-algebra containing a topology τ as a generator is called a *Borel σ-algebra* and it is denoted $\mathcal{B}(\tau)$ or by $\mathcal{B}(\Omega)$ or just $\mathcal{B}$ whenever the nature of τ is specified. Of the various Borel σ-algebras we come across, some are generated by the usual topology. □

2.3 Example. Let $\mathcal{G}$ be the system of sets containing only one subset A of Ω. As mentioned in Example 1.2(iii), the smallest σ-algebra containing $\mathcal{G}$ is $\{\Omega, \emptyset, A, A^c\}$. □

Problem 1.10 states that a Dynkin system is a σ-algebra if and only if it is $\cap$-stable. The proposition below generalizes it by allowing the Dynkin system to have just a $\cap$-stable generator.

2.4 Proposition. *A Dynkin system is a σ-algebra if and only if it has an $\cap$-stable generator.*

Proof. Let $\mathcal{G}$ be an $\cap$-stable system of subsets of Ω. Then $\Sigma(\mathcal{G}) = \mathcal{D}(\mathcal{G})$. Because every σ-algebra is a Dynkin system and $\mathcal{D}(\mathcal{G})$ is the smallest Dynkin system containing $\mathcal{G}$, $\mathcal{D}(\mathcal{G}) \subseteq \Sigma(\mathcal{G})$. The inverse relation remains to be shown:

Let $D \in \mathcal{D}(\mathcal{G})$ and let

$$\mathcal{D}_{\mathrm{D}} = \{Q \in \mathcal{P}(\Omega) : Q \cap D \in \mathcal{D}(\mathcal{G})\}.$$

$a)$ We show that $\mathcal{D}_D$ is a Dynkin system.

If $A \in \mathcal{D}_D$ then $A \cap D \in \mathcal{D}(\mathcal{G})$ and

$$A^c \cap D = D \backslash A = D \backslash (A \cap D) \in \mathcal{D}(\mathcal{G})$$

(see Problem 1.9). This yields $A^c \in \mathcal{D}_D$. Similarly, let $\{A_n\} \subseteq \mathcal{D}_D$ be a sequence of pairwise disjoint sets. Then

$$A_n \cap D \in \mathcal{D}(\mathcal{G}), \text{ for } n = 1, 2, \ldots.$$

Obviously, $\{A_n \cap D\}$ is a sequence of pairwise disjoint sets and

$$\sum_{n=1}^{\infty} A_n \cap D = \left(\sum_{n=1}^{\infty} A_n\right) \cap D \in \mathcal{D}(\mathcal{G}),$$

implying that $\sum_{n=1}^{\infty} A_n \in \mathcal{D}_D$. Therefore, $\mathcal{D}_D$ is a Dynkin system.

$b)$ We prove that for every $D \in \mathcal{D}(\mathcal{G})$, $\mathcal{D}(\mathcal{G}) \subseteq \mathcal{D}_D$.

Let $G \in \mathcal{G}$. Then $G \in \mathcal{D}(\mathcal{G})$. In as much as $\mathcal{G}$ is $\cap$-stable, it follows directly from the definition of $\mathcal{D}_G$ that $\mathcal{G} \subseteq \mathcal{D}_G$. Thus $\mathcal{G} \subseteq \mathcal{D}(\mathcal{G}) \subseteq \mathcal{D}_G$, since $\mathcal{D}(\mathcal{G})$ is the smallest Dynkin system containing $\mathcal{G}$ and $\mathcal{D}_G$ is just a Dynkin system containing $\mathcal{G}$. Now let $D \in \mathcal{D}(\mathcal{G})$. Then $D \in \mathcal{D}_G$ and $G \cap D \in \mathcal{D}(\mathcal{G})$, implying that $G \in \mathcal{D}_D$ or $\mathcal{G} \subseteq \mathcal{D}_D$. This yields that $\mathcal{D}(\mathcal{G}) \subseteq \mathcal{D}_D$, because again $\mathcal{D}(\mathcal{G})$ is the smallest Dynkin system containing $\mathcal{G}$.

c) We show that $\mathcal{D}(\mathcal{G})$ is $\cap$-stable.

Let $C, D \in \mathcal{D}(\mathcal{G})$. Then $\mathcal{D}(\mathcal{G}) \subseteq \mathcal{D}_D$ and $C, D \in \mathcal{D}_D$. Thus, $C \cap D \in \mathcal{D}(\mathcal{G})$ (by the definition of $\mathcal{D}_D$), and therefore $\mathcal{D}(\mathcal{G})$ is $\cap$-stable. Finally, by Problem 1.10, $\mathcal{D}(\mathcal{G})$ must be a σ-algebra. Then, as the smallest σ-algebra, $\Sigma(\mathcal{G}) \subseteq \mathcal{D}(\mathcal{G})$. This is the desired inverse relation. □

In the next lemma and theorem we present a construction of the ring generated by a semi-ring.

2.5 Lemma. *Let $\mathcal{S}$ be a semi-ring in Ω and let $\mathcal{R}$ be the system of all finite unions of elements from $\mathcal{S}$. Then any element of $\mathcal{R}$ can be represented as a finite union of pairwise disjoint sets from $\mathcal{S}$; in notation, $\sum_{k=1}^{n} C_k$, $C_k \in \mathcal{S}$.*

Proof. Let $R \in \mathcal{R}$. Then by the definition of $\mathcal{R}$, $R = \bigcup_{k=1}^{n} S_k$, where $S_k \in \mathcal{S}$. We now construct a decomposition of R by elements of $\mathcal{S}$ using S_k. Let

$$R_1 = S_1,\ R_2 = S_2 \backslash S_1,\ R_3 = [S_3 \backslash S_1] \cap [S_3 \backslash S_2],\ \dots,$$

$$R_k = \bigcap_{j=1}^{k-1} [S_k \backslash S_j],\ k = 1, \dots, n.$$

Because $S_k \backslash S_j = \sum_i C_{ikj}$ is a finite union of elements from $\mathcal{S}$, it follows that

$$R_k = \bigcap_{j=1}^{k-1} \sum_i C_{ikj} = \sum_i \bigcap_{j=1}^{k-1} C_{ikj}$$

(as a finite union of finite intersections)

$$= \sum_i D_{ik},$$

where D_{ik} are elements of $\mathcal{S}$. It is easily seen that

$$R = \bigcup_{k=1}^{n} S_k = \bigcup_{k=1}^{n} R_k,$$

where R_k are pairwise disjoint. This leads to $R = \sum_{k=1}^{n} \sum_i D_{ik}$ and the lemma is proved. □

2.6 Theorem. *Let $\mathcal{S}$ be a semi-ring in Ω. Then the system of all finite unions of sets in $\mathcal{S}$ is the (smallest) ring $\mathcal{R}(\mathcal{S})$ generated by $\mathcal{S}$.*

Proof.

1) We show that $\mathcal{R}$ described above is a ring. Because $\mathcal{S} \subseteq \mathcal{R}$, we have $\emptyset \in \mathcal{R}$. Let $R_1, R_2 \in \mathcal{R}$. Then, by the definition of $\mathcal{R}$,

$$R_1 = \bigcup_{k=1}^{n} S_k^1,\ R_2 = \bigcup_{i=1}^{m} S_i^2.$$

Therefore,

$$R_1 \cup R_2 = \bigcup_{k=1}^{n} \bigcup_{i=1}^{m} (S_k^1 \cup S_i^2) \in \mathcal{R}.$$

By Lemma 2.4 and by Problem 1.2(c) (Chapter 1), we have

$$R_1 \backslash R_2 = \left(\sum_{k=1}^{n} C_k\right) \backslash \left(\sum_{i=1}^{m} D_i\right)$$

$$= \bigcap_{i=1}^{m} \left(\sum_{k=1}^{n} (C_k \backslash D_i)\right) = \sum_{k=1}^{n} \bigcap_{i=1}^{m} C_k \backslash D_i.$$

Because C_k and D_i, as elements of $\mathcal{S}$, are semi-ring sets, the sets $C_k \backslash D_i = \sum_{s=1}^{r_{ik}} E_s^{ik}$ and E_s^{ik} are also elements of $\mathcal{S}$. Therefore,

$$R_1 \backslash R_2 = \sum_{k=1}^{n} \bigcap_{i=1}^{m} \sum_{s=1}^{r_{ik}} E_s^{ik} = \sum_{k=1}^{n} \sum_{s=1}^{r_{ik}} \bigcap_{i=1}^{m} E_s^{ik} \in \mathcal{R}.$$

2) Now we show that $\mathcal{R} = \mathcal{R}(\mathcal{S})$. Let $\mathcal{S} \subseteq \mathcal{R}'$, where $\mathcal{R}'$ is any ring in Ω. As a ring, $\mathcal{R}'$ is closed with respect to the formation of

finite unions of sets from $\mathcal{R}'$. Specifically, it is closed under finite unions of sets from $\mathcal{S}$; hence, it includes $\mathcal{R}$. Consequently, $\mathcal{R}$ is the smallest ring generated by $\mathcal{S}$. □

In Remark 2.2, we defined a Borel σ-algebra as a σ-algebra generated by a topology. We show below that the smallest σ-algebra $\Sigma(\mathcal{S}(\mathbb{R}^n))$ generated by the semi-ring of all half-open intervals $(a, b]$ in $\mathbb{R}^n$ coincides with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$.

2.7 Theorem. *Let τ_e be the usual topology on $\mathbb{R}^n$ and τ_e^c, and $\mathcal{K}_e$ denote, respectively, the systems of all closed and compact subsets of $\mathbb{R}^n$. Then the following relations hold.*

$$\mathcal{B}(\mathbb{R}^n) : = \Sigma(\tau_e) = \Sigma(\tau_e^c) = \Sigma(\mathcal{K}_e) = \Sigma(\mathcal{S}(\mathbb{R}^n)) .$$

Proof. Because all compact sets in $(\mathbb{R}^n, \tau_e)$ are closed and bounded, it follows that

$$\mathcal{K}_e \subseteq \tau_e^c \subseteq \Sigma(\tau_e^c),$$

and thus

$$\Sigma(\mathcal{K}_e) \subseteq \Sigma(\tau_e^c). \tag{2.7}$$

On the other hand, every closed set F can be represented as a countable union of compact sets $C_k \in \tau_e^c,\ k = 1, 2, \dots$. For instance, if $C(c, k)$ denotes the compact ball centered at some point c and with radius $k \in \mathbb{N},$ then we may choose $C_k = F \cap C(c, k)$ implying that $F = \bigcup\limits_{k=1}^{\infty} C_k$. Therefore, all closed sets belong to the σ-algebra $\Sigma(\mathcal{K}_e)$ (because this σ-algebra contains countable unions); that is, $\tau^c \subseteq \Sigma(\mathcal{K}_e)$ which yields

$$\Sigma(\tau_e^c) \subseteq \Sigma(\mathcal{K}_e). \tag{2.7a}$$

Both inclusions (2.7) and (2.7a) lead to $\Sigma(\tau_e^c) = \Sigma(\mathcal{K}_e)$. Because open sets are complementary to closed sets, it follows that

$$\mathcal{B} = \Sigma(\tau_e) = \Sigma(\tau_e^c) = \Sigma(\mathcal{K}_e).$$

Now we show that $\mathcal{B} = \Sigma(\mathcal{S})$. Any half-open interval $(a, b]$ in $\mathbb{R}^n$ can be represented as the intersection of a sequence of bounded open intervals of type (a, b_n) (or as we called them earlier, *open rectangles*) such that $b_n \downarrow b$. Therefore, the collection $\mathcal{S}$ of all half-open intervals belongs to the σ-algebra $\Sigma(\tau_e)$, which implies that $\Sigma(\mathcal{S}) \subseteq \Sigma(\tau_e)$. On the other hand, any open bounded interval can be represented as the union of a sequence of half-open intervals of $\mathcal{S}$; and any open set is a countable union of bounded open intervals as base sets (recall that $(\mathbb{R}^n, \tau_e)$ is second countable). Therefore, any open set is the union of countably many half-open intervals from $\mathcal{S}$ and we have $\tau_e \subseteq \Sigma(\mathcal{S})$, implying $\Sigma(\tau_e) \subseteq \Sigma(\mathcal{S})$. Dual containment yields $\Sigma(\mathcal{S}) = \mathcal{B}$. □

PROBLEMS

2.1 Prove Theorem 2.1. *The intersection of arbitrarily many σ-algebras* (*algebras*, *monotone systems*, *rings*) *in Ω is a σ-algebra* (*an algebra*, *a monotone system*, *a ring*).

2.2 Show that an intersection of semi-rings in Ω need not be a semi-ring.

2.3 Show that a union of σ-algebras in Ω need not be a σ-algebra.

2.4 Let A and B be nonempty subsets of Ω and let $\mathcal{G} = \{A, B\}$. Find $\mathcal{D}(\mathcal{G})$ and $\Sigma(\mathcal{G})$. Show that $\mathcal{D}(\mathcal{G})$ and $\Sigma(\mathcal{G})$ are identical if and only if one of the following conditions holds.

$A \cap B$ or $A \cap B^c$ or $A^c \cap B$ or $A^c \cap B^c$ is the empty set.

(a) Use Problem 1.10 to justify it.

(b) Use Proposition 2.4 to justify it.

2.5 Let Σ be a σ-algebra in Ω and let $B \subseteq \Omega$. Show that the σ-algebra generated by $\mathcal{G} = \Sigma \cup \{B\}$ is of the form

$$\Sigma' = \{(A_1 \cap B) \cup (A_2 \backslash B) \colon A_1, A_2 \in \Sigma\}.$$

[*Hint*: 1) Show that $\sigma(\Sigma') = \sigma(\Sigma \cup \{B\})$. 2) Show that Σ' is a σ-algebra in Ω.]

2.6 Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be systems of sets in Ω. Show that $\mathcal{G}_1 \subseteq \mathcal{G}_2$ implies that $\Sigma(\mathcal{G}_1) \subseteq \Sigma(\mathcal{G}_2)$.

2.7 Let Ω be an arbitrary nonempty set and let $A, B \subseteq \Omega$. Construct for the topology $\tau = \{\Omega, \emptyset, A, B, A \cap B, A \cup B\}$ the Borel σ-algebra $\mathcal{B}(\tau)$.

2.8 Construct the Borel σ-algebra generated by the cocountable topology $\tau = \{\Omega, \emptyset, A^c : A \text{ is at most countable}\}$ (see Problem 1.9, Chapter 3, where Ω is an uncountable set).

2.9 Show that any open set in $\mathbb{R}^n$ can be represented as at most a countable union of disjoint half-open cubes.

NEW TERMS:

σ-algebra generated by a collection of sets, 247
generator, 247
$\mathcal{B}(\tau)$, Borel σ-algebra, 247
σ-algebra generated by a set, 248
semi-ring, property of, 249
ring generated by a semi-ring, 250
$\mathcal{B}(\mathbb{R}^n)$, Borel σ-algebra in Euclidean space, 251
Borel σ-algebra generated by a cocountable topology, 253

3. MEASURABLE FUNCTIONS

Measurable functions which we introduce in this section are analogues of continuous functions f operating between two topological spaces, say (X, τ) and (Y, τ'). Recall that continuity of f was first defined at a point and then on X followed by a theorem that f was continuous if and only if $f^{**}(\tau') \subseteq \tau$. In the present case, a function acting between two measurable spaces (Ω, Σ) and (Ω', Σ') is regarded as *measurable* if $f^{**}(\Sigma') \subseteq \Sigma$.

3.1 Definition. Let (Ω, Σ) and (Ω', Σ') be two measurable spaces. A function $[\Omega, \Omega', f]$ is called *measurable* if $f^{**}(\Sigma') \subseteq \Sigma$, that is, if $\forall A' \in \Sigma'\ f^*(A') \in \Sigma$. The collections of all measurable functions from (Ω, Σ) to (Ω', Σ') are denoted by the symbol $\mathcal{C}^{-1}(\Omega, \Sigma; \Omega', \Sigma')$.

Notice that symbol $\mathcal{C}^{-1}$ is a natural extrapolation of the common notation in analysis, where $\mathcal{C}^n$ stands for the space of all n-times continuously differentiable functions, with $\mathcal{C}^0$ (or simply $\mathcal{C}$) standing for the space of all continuous functions. So, not only has $\mathcal{C}^{-1}$ been vacant, but it also agrees with the existing linear order ($\mathcal{C}^n, n = -1, 0, 1, \ldots; \supseteq$). □

3.2 Remark. In Example 1.2$(viii)$, we saw that $f^{**}(\Sigma')$ is a σ-algebra in Ω. We wish to call it the *σ-algebra generated by function f*. This is the smallest σ-algebra relative to which f is measurable. Recall that if τ' is a topology on Ω', then $f^{**}(\tau')$ is the weakest topology on Ω relative to which f is continuous. □

3.3 Examples.

(i) Each identity function $f\colon (\Omega, \Sigma) \to (\Omega, \Sigma')$ is measurable if and only if $\Sigma' \subseteq \Sigma$.

(ii) Let $f\colon (\Omega, \Sigma) \to (\Omega', \Sigma')$ be a constant function, that is, $f(\omega) = c \in \Omega',\ \forall \omega \in \Omega$. Therefore, $f^*(\{c\}) = \Omega$ and $f^*(\{c\}^c) = \emptyset$, which yields that for each $A' \in \Sigma,\ f^*(A')$ is either Ω or $\emptyset$. The latter implies that f is measurable with respect to the smallest σ-algebra $\{\Omega, \emptyset\}$ in Ω. Thus, f is always measurable.

(iii) Let $f = \mathbf{1}_A$ for some $A \subseteq \Omega$. Let Σ' be an arbitrary σ-algebra in $\mathbb{R}$ (for instance, the Borel σ-algebra). It is easily seen that the inverse image under f of any subset of $\mathbb{R}$ (specifically, of any subset of Σ') is one of the elements of the set $\Sigma = \{\Omega, \emptyset, A, A^c\}$. Therefore, Σ is the smallest σ-algebra with respect to which f is measurable. On the other hand, if Σ is a σ-algebra in Ω, then $\mathbf{1}_A$ is measurable if and only if $A \in \Sigma$. □

There is a noteworthy parallel between continuity and measurability of functions and their relationships with topologies and σ-algebras mentioned in the beginning of this section. Recall that a function $[\Omega, \Omega', f]$ is continuous on Ω if there are two topologies τ and τ' declared on Ω and Ω', respectively, and $f^{**}(\tau') \subseteq \tau$. If, in addition, τ' is known to be induced by a subbase $\mathcal{G}'$, then the condition $f^{**}(\tau') \subseteq \tau$ can be relaxed by $f^{**}(\mathcal{G}') \subseteq \tau$, as per Theorem 4.8, Chapter 3. The pointed-out analogue of continuity with measurability is utilized by the following proposition.

3.4 Proposition. *Let (Ω, Σ) and (Ω', Σ') be two measurable spaces and let $\mathcal{G}'$ be a generator of Σ'. Then a function $f\colon \Omega \to \Omega'$ is measurable if and only if $f^{**}(\mathcal{G}') \subseteq \Sigma$.*

Proof. Let

$$\widetilde{\Sigma} = \{Q' \subseteq \Omega' : f^*(Q') \in \Sigma\}. \tag{3.4}$$

It can be shown that $\widetilde{\Sigma}$ is a σ-algebra (Problem 3.1), which resembles the quotient topology. Now let $f^{**}(\mathcal{G}') \subseteq \Sigma$. Then from (3.4) it follows that $\mathcal{G}' \subseteq \widetilde{\Sigma}$ and hence by Problem 2.6,

$$\Sigma' = \Sigma(\mathcal{G}') \subseteq \widetilde{\Sigma}.$$

Therefore, f is measurable. Obviously, $\widetilde{\Sigma}$ is largest σ-algebra with respect to which function f is measurable. The converse is trivial. □

3.5 Examples.

(i) Let $f\colon (\Omega, \tau) \to (\Omega', \tau')$ be a continuous function on a topological space (Ω, τ). Then $f^{**}(\tau') \subseteq \tau \subseteq \mathcal{B}(\tau)$ (the Borel σ-algebra generated by τ). By Proposition 3.4, the function f is then

$\mathcal{B}(\tau)$-$\mathcal{B}(\tau')$-measurable. Here f is an example of a *Borel measurable function.*

(ii) In a more general setting, $f : (\Omega, \mathcal{B}(\tau)) \to (\Omega', \mathcal{B}(\tau'))$ is called a *Borel measurable function* or just *Borel function*, if

$$f^{**}(\mathcal{B}(\tau')) \subseteq \mathcal{B}(\tau).$$

Now, by Proposition 3.4, f is Borel if and only if $f^{**}(\tau') \subseteq \mathcal{B}(\tau)$, that is, if the inverse image of an open set is a Borel set.

(iii) Let $f : (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ be a function. Because by Theorem 2.7, the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ has a variety of generators, including $\tau_e, \tau_e^c, \mathcal{K}_e, \mathcal{S}$, to see whether f is Borel, we can examine if $f^{**}(\,\cdot\,)$ lands at $\mathcal{B}(\mathbb{R}^n)$ with one of these families. The most common family used for verification is the semi-ring $\mathcal{S}$ of half-open rectangles.

For example, let $f = \mathbf{1}_{\mathbb{Q}}$ (the Dirichlet function). Obviously, $\mathbb{Q}$ is a Borel set, because $\mathbb{Q}$ is the countable union of the singletons representing all rational points. Each one of them is a closed set, because τ_e is Hausdorff. If $(a, b] \in \mathcal{S}$ $(a < b)$, $f^*((a, b])$ is either $\emptyset$, or $\mathbb{Q}$, or $\mathbb{Q}^c$, or $\mathbb{R}$. Hence the Dirichlet function is Borel. □

Measurability, like continuity, is preserved under the composition.

3.6 Proposition. *Let* $f_1 \colon (\Omega_1, \Sigma_1) \to (\Omega_2 = f_{1*}(\Omega_1), \Sigma_2)$ *and* $f_2 \colon (\Omega_2, \Sigma_2) \to (\Omega_3, \Sigma_3)$ *be measurable functions. Then the composition* $f_2 \circ f_1 : (\Omega_1, \Sigma_1) \to (\Omega_3, \Sigma_3)$ *is measurable.* □

(See Problem 3.2.)

3.7 Remark. Let $\{\Omega_i, \Sigma_i \colon i \in I\}$ be an arbitrary collection of measurable spaces and let $\{f_i \colon \Omega \to \Omega_i : i \in I\}$ be a collection of functions defined on a set Ω. Every function f_i of this family is clearly $f_i^{**}(\Sigma_i)$-Σ_i-measurable. We are interested in constructing the minimal σ-algebra in Ω relative to which all functions of the family are measurable. Since $\bigcup_{i \in I} f_i^{**}(\Sigma_i)$ is not, in general, a σ-algebra, it is reasonable to regard it as the generator of the σ-algebra *generated by*

the family $\{f_i;\ i \in I\}$, in notation, $\Sigma(f_i; i \in I)$. Recall the analogue with continuous functions of Remark 5.10, Chapter 3. □

3.8 Lemma. *Let* $\{g_i\colon (\Omega, \Sigma) \to (\Omega_i, \Sigma_i)\}$ *be a collection of functions on* Ω *and let* $f:\ (\Omega_0, \Sigma_0) \to (\Omega, \Sigma)$ *be a function on* Ω_0. *The function* f *is* Σ_0-$\Sigma(g_i\colon i \in I)$*-measurable if and only if each of the functions* $g_i \circ f$ is Σ_0-Σ_i*-measurable*.

Proof.

1) Let $g_k \circ f$ be Σ_0-Σ_k-measurable $\forall k \in I$. Then $\forall A_k \in \Sigma_k$,

$$(g_k \circ f)^*(A_k) = f^* \circ g_k^*(A_k) \in \Sigma_0,$$

where $g_k^*(A_k) \in \bigcup\limits_{i\in I} g_i^*(\Sigma_i)$. Taking A_k from Σ_k for each respective $k \in I$ we run the whole set $\bigcup\limits_{i\in I} g_i^{**}(\Sigma_i)$ whose elements are further transferred by f^* into Σ_0. In other words, we have

$$f^{**}(\bigcup_{i\in I} g_i^*(\Sigma_i)) \subseteq \Sigma_0. \tag{3.8}$$

Because $\bigcup\limits_{i\in I} g_i^{**}(\Sigma_i)$ is a generator of $\Sigma(g_i;\ i \in I)$, according to Proposition 3.4, inclusion (3.8) is sufficient for $f^{**}(\Sigma(g_i; i \in I)) \subseteq \Sigma_0$. Therefore, f is indeed Σ_0-$\Sigma(g_i; i \in I)$-measurable.

2) Let f be Σ_0-$\Sigma(g_i; i \in I)$-measurable. This implies that

$$\forall E \in \mathcal{G} = \bigcup_{i\in I} g_i^{**}(\Sigma_i),\ f^*(E) \in \Sigma_0.$$

Besides, $\forall A_k \in \Sigma_k$,

$$g_k^*(A_k) \in \bigcup_{i\in I} g_i^{**}(\Sigma_i).$$

Thus, $\forall A_k \in \Sigma_k,\ \ (g_k \circ f)^*(A_k) \in \Sigma_0,$ which means that $g_k \circ f$ is Σ_0-Σ_k-measurable. □

PROBLEMS

3.1 Show that $\widetilde{\Sigma}$ in the proof of Proposition 3.4 is a σ-algebra.

3.2 Prove Proposition 3.6: *Let* $f_1\colon (\Omega_1, \Sigma_1) \to (\Omega_2 = f_{1*}(\Omega_1), \Sigma_2)$ *and* $f_2\colon (\Omega_2, \Sigma_2) \to (\Omega_3, \Sigma_3)$ *be measurable functions. Then the composition* $f_2 \circ f_1 : (\Omega_1, \Sigma_1) \to (\Omega_3, \Sigma_3)$ *is measurable.*

3.3 Let $f : \Omega \to \Omega'$ be a function and let $\mathcal{G}' \subseteq \mathcal{P}(\Omega')$. Show that

$$f^{**}(\Sigma(\mathcal{G}')) = \Sigma(f^{**}(\mathcal{G}')).$$

[*Hint*: Let $\widetilde{\Sigma}\colon = \{A' \in \mathcal{P}(\Omega')\colon\ f^*(A') \in \Sigma(f^{**}(\mathcal{G}'))\}$. Show that $\widetilde{\Sigma}$ is a σ-algebra. Then show that $\Sigma(\mathcal{G}') \subseteq \widetilde{\Sigma}$.]

3.4 Let $[O, O_1, F]$ be a homeomorphism, with O and O_1 being open sets in topological spaces (X, τ) and (X, τ_1), respectively, and let $\mathcal{B}(\tau_O)$ and $\mathcal{B}(\tau_{O_1})$ be the Borel σ-algebras generated by the relative topologies τ_O and τ_{O_1}. Prove that $[\mathcal{B}(\tau_O), \mathcal{B}(\tau_{O_1}), F_*]$ is bijective.

3.5 Let $[O, O_1, F]$ be a homeomorphism, with O and O_1 being open sets in topological spaces (X, τ) and (X, τ_1), respectively, and let $\mathcal{B}(\tau_O)$ and $\mathcal{B}(\tau_{O_1})$ be the corresponding Borel σ-algebras generated by the relative topologies τ_O and τ_{O_1}. Suppose $B \subseteq O$. Show that if $F_*(B)$ is Borel, then B is also Borel.

NEW TERMS:

measurable function, 255
$\mathcal{C}^{-1}(\Omega, \Sigma; \Omega', \Sigma')$, set of measurable functions, 255
σ-algebra generated by a function, 255
measurability of a function, criterion of, 256
Borel measurable function, 257
composition of measurable functions, 257
σ-algebra generated by a family of functions, 257
homeomorphisms and Borel σ-algebras, relationship between, 259

Chapter 5

Measures

This chapter is a precursor to the general theory of integration, which is a significant advancement from the Riemann integration known from calculus. Although many applications in the natural sciences triggered the development of general integration and measure theory, the theory of probability has become the primary client of abstract measure even prior to integration.

An early notion of measure was introduced by the Italian Giuseppe Peano (1858 - 1932) in 1883. For a simple set in the plane, Peano used two types of polygons that contain and are included in the given set. The areas of the polygons of the former type have a greatest lower bound and of the latter type, the least upper bound. If these limits coincide, their common value is said to be the *area of the set*. However, if the limits differ, the concept of area would not apply. Apparently, Cantor's development of set theory greatly influenced Peano's concept of area for arbitrary sets in his 1887 monograph, *Applicazioni geometriche del calcolo infinitesimale*. He generalized his original idea on inner and outer measures of sets by polygons for two- and three-dimensional Euclidean sets. Peano emphasized the close connection between measure and integral.

In 1892, the Frenchman Camille Jordan (1838 - 1922) arrived at a more advanced concept of measure as a countably additive set function applied first to positive and then to signed set functions. The latter led to the prominent Jordan decomposition of a signed measure in two positive, which we will study in Chapter 6, section 6. Jordan's motivation of the concept of measurable sets apparently stemmed

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0_5, © Springer Science+Business Media, LLC 2013

from the theory of double integration, which naturally arises when introducing integrals on arbitrary plane sets. However, Jordan's approach to the measurements of sets was restricted to the common, at that time, finite covers of sets by intervals or rectangles.

The most revolutionary steps were undertaken by the Frenchman Émile Borel (1871 - 1956) in his famous 1898 monograph, *Leçons sur la théorie des fonctions*, where he introduced the idea of countable, instead of finite, covers, thereby significantly extending classes of measurable sets. Borel also pointed out in 1905 to a possibility of using measure theory in probability, which was successfully accomplished by the Russian Andrey Kolmogorov (1903 - 1987) no earlier than 1933. However, in his *Leçons*, Borel did not bother to connect measure and integration.

In 1902, another Frenchman, Henri Lebesgue (1875 - 1941), further refined measure theory by combining the ideas of Camille Jordan on finite contents with the countably additive measure notion of Émile Borel. Lebesgue called sets in $\mathbb{R}^n$ *measurable* whenever their inner and outer measures are equal. This led to the completion of the concept of measure and gave rise to the general theory of integration so significantly enlarging the class of integrable functions that it made Lebesgue say in 1902: “I know no function that is not summable and I do not know if any such exists.” (However, the Italian Giuseppe Vitali showed the existence of such a function in 1905.) Lebesgue also established several central theorems in the theory of integration; one of them is the famous Lebesgue dominated convergence theorem.

Finally, the Austrian Johann Radon (1887-1956), in his 1913 habilitation work began to study abstract measures and integrals more general than those of Lebesgue. Radon is also the author of the well-known Radon and Radon - Stieltjes integrals. The latter is most frequently referred to as the *Lebesgue - Stieltjes integral.* Radon's ideas led not only to the abstract theory of measure and integration but also to its applications in the boundary value problems in the theory of logarithmic potentials (developed by Radon himself) and contemporary theory of probability and stochastic processes.

1. SET FUNCTIONS

1.1 Definitions.

(i) Let Σ be a system of subsets of Ω including the empty set $\emptyset$. A numerical function $[\Sigma, \overline{\mathbb{R}}, \mu]$ such that $\mu(\emptyset) = 0$ is called a *set function*. In this chapter we only consider nonnegative set functions $\mu : \Sigma \to \overline{\mathbb{R}}_+$.

In the definitions below we assume that corresponding sets are elements of Σ.

(ii) A set function μ is called *finitely additive* or just *additive* if for any n-tuple $A_1, \ldots, A_n$ of mutually disjoint sets with $\sum_{k=1}^{n} A_k \in \Sigma$,

$$\mu\left(\sum_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k).$$

(iii) A set function μ is called *σ-additive* if for any sequence $A_1, A_2, \ldots$ of mutually disjoint sets with $\sum_{n=1}^{\infty} A_n \in \Sigma$, it holds that

$$\mu\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

(iv) A set function μ on Σ is called *continuous from below on* Σ if for every monotone nondecreasing sequence $\{A_n\}\uparrow$ such that $\bigcup_{n=1}^{\infty} A_n \in \Sigma$ it holds that

$$\lim_{n \to \infty} \mu(A_n) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right).$$

If this condition is known to hold just on a particular monotone nondecreasing sequence $\{A_n\}\uparrow$, then μ is said to be *continuous from below on* $\{A_n\}$.

(v) Let $\{A_n\}$ be a monotone nonincreasing sequence and $\bigcap_{n=1}^{\infty} A_n \in \Sigma$. A set function μ on Σ is said to be *continuous from above on* $\{A_n\}$ if

$$\lim_{n \to \infty} \mu(A_n) = \mu\left(\bigcap_{n=1}^{\infty} A_n\right). \tag{1.1}$$

The set function μ is called *continuous from above on* Σ if equation (1.1) holds for every monotone nonincreasing sequence $\{A_n\}\downarrow$ such that $\bigcap_{n=1}^{\infty} A_n \in \Sigma$. In particular, if $\{A_n\}\downarrow \emptyset$, (1.1) reduces to

$$\lim_{n\to\infty} \mu(A_n) = 0$$

and this is referred to as *continuity from above at the empty set* or, for short, $\emptyset$-*continuity* of μ.

(vi) A set function μ is called *finite* on Σ if $\mu(A) < \infty$ for all $A \in \Sigma$.

(vii) A set function μ is called σ-*finite* on Σ if Σ contains a sequence $\{A_n\}$ monotonically increasing to Ω (that is, $\bigcup_{n=1}^{\infty} A_n = \Omega$) and μ is finite on $\{A_n\}$. In this case, we also say that μ is σ-*finite on* $\{A_n\}$.

$(viii)$ An additive set function μ defined on a semi-ring $\mathcal{S}$ (in Ω) is called an *elementary content* (*on* $\mathcal{S}$).

(ix) An additive set function μ defined on a ring $\mathcal{R}$ (in Ω) is called a *content* (*on* $\mathcal{R}$).

(x) A σ-additive set function μ defined on a ring $\mathcal{R}$ or algebra $\mathcal{A}$ (in Ω) is called a *premeasure* (*on* $\mathcal{R}$ or $\mathcal{A}$).

(xi) A σ-additive set function μ defined on a σ-algebra Σ (in Ω) is called a *measure* (*on* Σ). If, in addition, $\mu(\Omega) = 1$, then μ is called a *probability measure.*

(xii) Let (Ω, Σ) be a measurable space and let μ be a measure on Σ. Then the triple (Ω, Σ, μ) is called a *measure space.* If μ is a probability measure, then the triple (Ω, Σ, μ) is called a *probability space.* □

1.2 Examples.

(i) Let Σ be a σ-algebra in Ω and let $a \in \Omega$ be a fixed point. Define the following set function ε_a on Σ. For each $A \in \Sigma$, we set $\varepsilon_a(A) = 1$ if $a \in A$ and $\varepsilon_a(A) = 0$ if $a \notin A$. Then ε_a is a measure on Σ. It is clear that $\varepsilon_a(\emptyset) = 0$ and that $\varepsilon_a \geq 0$. Let $\{A_n\}$ be a

sequence of pairwise disjoint sets from Σ. Then a can either belong to exactly one set, A_j (for some j), or to no set of the sequence. In the first case,

$$\varepsilon_a\left(\sum_{n=1}^{\infty} A_n\right) = 1,$$

and in the second case,

$$\varepsilon_a\left(\sum_{n=1}^{\infty} A_n\right) = 0.$$

On the other hand, in the first case,

$$\sum_{n=1}^{\infty} \varepsilon_a(A_n) = 1,$$

and in the second case,

$$\sum_{n=1}^{\infty} \varepsilon_a(A_n) = 0.$$

Therefore, ε_a is σ-additive. The measure ε_a is called a *point mass* or *Dirac measure*.

(ii) Let $\Omega = \mathbb{R}^n$. In Example 1.2(iv), Chapter 4, we introduced the system $\mathcal{S}$ of all half-open intervals $(a, b] \subseteq \mathbb{R}^n$, which was shown to be a semi-ring on $\mathbb{R}^n$. For $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ $(a_i < b_i)$, we define

$$\lambda^0((a, b]) = \prod_{k=1}^{n} (b_k - a_k) \text{ and } \lambda^0(\emptyset) = 0.$$

Then λ^0 is obviously an elementary content on $\mathcal{S}$. λ^0 is said to be the *Lebesgue elementary content*.

(iii) Let $[\mathbb{R}, \mathbb{R}, f]$ be a bounded, monotone nondecreasing, right-continuous function that vanishes at $-\infty$. Any such function is called a *distribution function*. There is also a variant of the so-called *extended distribution function* $[\mathbb{R}, \mathbb{R}, f]$, which is monotone nondecreasing, right-continuous, but need not be bounded over unbounded sets and does not vanish at $-\infty$. (As a right-continuous function, an extended distribution function is bounded over bounded sets though.) Because f is monotone nondecreasing and right continuous, for any two real numbers $a < b$, the set $f^*((a, b])$ is Borel (show it; Problem 1.15). In other words, we have $f^{**}(\mathcal{S}) \subseteq \mathcal{B}(\mathbb{R})$, where $\mathcal{S}$ is the semi-

ring of left half-open intervals. Thus, by Proposition 3.4, Chapter 4 [see also Example 3.5 (iii), Chapter 4], f is Borel.

So, every distribution or extended distribution function is Borel. Both types of functions are subclasses of monotone functions (which are introduced and studied in forthcoming chapters).

Let $\mathcal{S}$ be the semi-ring of half-open intervals in $\mathbb{R}$. Given a distribution function f, we define the set function μ_f^0 on $\mathcal{S}$ as

$$\mu_f^0(\emptyset) = 0 \text{ and } \mu_f^0((a,b]) = f(b) - f(a).$$

μ_f^0 is clearly additive on $\mathcal{S}$ and therefore is an elementary content on $\mathcal{S}$.

μ_f^0 is called the *Lebesgue - Stieltjes elementary content*.

(iv) Let Ω be an uncountable set and let

$$\Sigma = \{A \in \mathcal{P}(\Omega) : \textit{either } A \textit{ or } A^c \preceq \mathbb{N}\},$$

which is a σ-algebra on Ω [see Example $1.2(vii)$, Chapter 4]. Then, define $\forall A \in \Sigma$,

$$\mu(A) = \begin{cases} 0, & A \preceq \mathbb{N} \\ 1, & A^c \preceq \mathbb{N} \end{cases}$$

We show that μ is a measure on Ω. First observe that $\mu \geq 0$ and $\mu(\emptyset) = 0$. Let $\{A_n : n = 1, 2, \dots\}$ be a sequence of pairwise disjoint sets of Σ. If the union $\Sigma_{n=1}^{\infty} A_n$ is at most countable, then each A_n is at most countable, and thus

$$\mu\left[\sum_{n=1}^{\infty} A_n\right] = 0 = \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} 0.$$

If $[\sum_{n=1}^{\infty} A_n]^c$ is at most countable, then we argue that there is exactly one set of the sequence $\{A_n\}$ with at most a countable complement. Suppose that there is yet another set from this sequence with this property. Then both of them, say A_i^c and A_k^c, are at most countable and hence $A_k^c \cap A_k^c$ is also at most countable. In as much

as $A_i \cap A_k = \emptyset$, $A_i^c \cup A_k^c = \Omega$, which is a contradiction, for Ω is uncountable. Therefore, we have exactly one set A_i such that A_i^c is at most countable. Then,

$$[\textstyle\sum_{n=1}^{\infty} A_n]^c = \bigcap_{n=1}^{\infty} A_n^c$$

is at most countable and

$$\mu\,[\textstyle\sum_{n=1}^{\infty} A_n] = 1.$$

On the other hand,

$$\textstyle\sum_{n=1}^{\infty} \mu(A_n) = \underbrace{\mu(A_1) + \cdots + \mu(A_{i-1})}_{0}$$

$$+ \underbrace{\mu(A_i)}_{1} + \underbrace{\mu(A_{i+1}) + \cdots}_{0} = 1.$$

This yields σ-additivity of μ.

(v) Let $\{\Omega, \Sigma, \mu_i : i = 1, 2, \dots\}$ be a sequence of measure spaces and let $\{\alpha_i : i = 1, 2, \dots\}$ be a nonnegative numerical sequence. Define μ on Σ as

$$\mu = \textstyle\sum_{n=1}^{\infty} \alpha_n \mu_n. \tag{1.2}$$

Then μ is a measure on Σ. (See Problem 1.4.)

(vi) Consider the special case in the last example with $\mu_n = \varepsilon_{b_n}$, $n = 1, 2, \dots,$ where $\{b_n\} \subseteq \Omega$. In other words, μ_n is a point mass that was introduced in (i). Then the measure μ defined in (1.2) is called an *atomic* (or *discrete*) *measure.* A further special case is of interest. With the given μ_n, we also assume that the sequence $\{\alpha_n\}$ has the property

$$\textstyle\sum_{n=1}^{\infty} \alpha_n = 1.$$

It is readily seen that the measure μ is a probability measure; more specifically, it is an *atomic* (or *discrete*) *probability measure.*

(vii) Let Ω be an uncountable set and let $\Sigma = \{A \in \mathcal{P}(\Omega)$: either A or $A^c \preceq \mathbb{N}\}$ as in (iv). Define $\forall A \in \Sigma,\ \mu(A) = 0$ if A is at most countable and $\mu(A) = \infty$ if A^c is at most countable. Then μ is a measure on Σ.

$(viii)$ Let Ω be an arbitrary set and Σ be a σ-algebra on Ω. Define the following set function μ on Σ. For each $A \in \Sigma,\ \mu(A) = |A|$ (that is, the number of elements in A), if A is finite and $\mu(A) = \infty$, otherwise. Then, it is easy to verify that μ is a measure on Σ. It is called a *counting measure* and (Ω, Σ, μ) is referred to as a *counting measure space*. Note that if Ω is at most a countable set, a counting measure μ can also be expressed in terms of the measure introduced in (1.2), with $\alpha_n = 1$ and $\mu_n = \varepsilon_{b_n}$. □

1.3 Proposition. *Let μ be an elementary content on a semi-ring $\mathcal{S}$. Then the following hold true.*

(i) *For any two sets A and B from $\mathcal{S}$ with $A \subseteq B,\ \mu(A) \leq \mu(B)$ (monotonicity).*

(ii) *Let $A_1, A_2, \ldots$ be a sequence of pairwise disjoint sets from $\mathcal{S}$ and $A_n \subseteq A \in \mathcal{S}$ for each n. Then*,

$$\sum_{n=1}^{\infty} \mu(A_n) \leq \mu(A). \tag{1.3}$$

(iii) *Let $A, A_1, A_2, \ldots \in \mathcal{S}$ and $A \subseteq \bigcup_{n=1}^{\infty} A_n$. Then there is a countable decomposition $\sum_{k=1}^{\infty} C_k$ of A with C_k's $\in \mathcal{S}$ such that*

$$\sum_{k=1}^{\infty} \mu(C_k) \leq \sum_{n=1}^{\infty} \mu(A_n). \tag{1.3a}$$

Proof.

(i) $B = A + B\backslash A = A + \sum_{s=1}^{n} C_s,$ where C_s's $\in \mathcal{S}$. Hence, by additivity of μ,

$$\mu(B) = \mu(A) + \sum_{s=1}^{n} \mu(C_s)$$

and the statement follows.

(ii) By Problem 1.12, Chapter 4,

$$A\backslash\textstyle\sum_{k=1}^{n} A_k = \sum_{s=1}^{N} C_s, \text{ where } C_s \in \mathcal{S}.$$

Hence, because of the assumption that $A_n \subseteq A$,

$$A = \textstyle\sum_{k=1}^{n} A_k + \sum_{s=1}^{N} C_s.$$

Thus, by additivity of μ on $\mathcal{S}$,

$$\mu(A) = \textstyle\sum_{k=1}^{n} \mu(A_k) + \sum_{s=1}^{N} \mu(C_s),$$

which implies that

$$\mu(A) \geq \textstyle\sum_{k=1}^{n} \mu(A_k), \text{ good for all } n.$$

Consequently,

$$\mu(A) \geq \textstyle\sum_{k=1}^{\infty} \mu(A_k).$$

With $B_n = A_n \cap A \ (\in \mathcal{S})$ we have that $\bigcup_{n=1}^{\infty} B_n = A$. Denote

$$D_1 = B_1,\ D_{n+1} = B_{n+1}\backslash \bigcup_{j=1}^{n} B_j,\ n = 1, 2, \ldots.$$

Then, $A = \sum_{k=1}^{\infty} D_k$ and, by Problem 1.12, Chapter 4,

$$D_k = \textstyle\sum_{s=1}^{j_k} C_{sk} \text{ with } C_{sk}\text{'s} \in \mathcal{S}.$$

Because $D_k \subseteq B_k$, by (ii),

$$\textstyle\sum_{s=1}^{j_k} \mu(C_{sk}) \leq \mu(B_k) \text{ for each } k.$$

Now, due to monotonicity of μ and because $B_k \subseteq A_k$, we have that

$$\textstyle\sum_{k=1}^{\infty} \sum_{s=1}^{j_k} \mu(C_{sk}) \leq \sum_{k=1}^{\infty} \mu(B_k) \leq \sum_{k=1}^{\infty} \mu(A_k), \qquad (1.3b)$$

which yields (1.3a). □

1.4 Corollaries.

(i) *If* $\sum_{m=1}^{\infty} E_m$ *is a decomposition of* A *by elements* E_m's *from* $\mathcal{S}$, *then* [from Proposition 1.3(iii)]:

$$\sum_{k=1}^{\infty} \mu(C_k) \leq \sum_{m=1}^{\infty} \mu(E_m); \tag{1.4}$$

i.e., $\sum_{k=1}^{\infty} C_k$ is a "μ-minimal decomposition" of A.

(ii) *Let* $A, A_1, A_2, \ldots \in \mathcal{S}$ *and* $A \subseteq \bigcup_{n=1}^{\infty} A_n$. *If* μ *is* σ*-additive on* $\mathcal{S}$, *that is*, *for any sequence of mutually disjoint sets* $\{A_n\} \subseteq \mathcal{S}$ *with* $\Sigma_{n=1}^{\infty} A_n \in \mathcal{S}$,

$$\mu(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n),$$

then (from (1.4)):

$$\mu(A) = \sum_{k=1}^{\infty} \sum_{s=1}^{j_k} \mu(C_{sk}) \leq \sum_{k=1}^{\infty} \mu(A_k). \tag{1.4a}$$

In particular, *if* $A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{S}$,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n), \tag{1.4b}$$

which is known as the σ-*subadditivity* property.

Inequality (1.4b) (originally applied to a semi-ring $\mathcal{S}$ and elementary content μ) implies:

(iii) *If* μ *is a content on a ring* $\mathcal{R}$ *then for any* $A_1, \ldots, A_n \in \mathcal{R}$,

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} \mu(A_k) \text{ (finite subadditivity).}$$

(iv) *If* μ *is a measure on a* σ*-algebra* Σ *then the* σ*-subadditivity property* (1.4b) *is valid for any sequence* $\{A_n\} \subseteq \Sigma$.

(v) *Because of the monotonicity property of an elementary content, due to Proposition* $1.3(i)$, *Definition* $(1.1(iv))$ *of finiteness can be relaxed for set functions from elementary contents and "above" by requiring merely* $\mu(\Omega) < \infty$. □

There are two more minor properties of contents left as an exercise.

1.5 Proposition. *Every content* μ *on a ring* $\mathcal{R}$ *has the following properties*.

(i) $A \subseteq B \Rightarrow \mu(B\backslash A) = \mu(B) - \mu(A)$ [*provided that* $\mu(A) < \infty$].

(ii) $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$. □

(See Problem 1.2.)

1.6 Lemma. *Let* μ *be a content on a ring* $\mathcal{R}$. *Then* μ *is a premeasure* (*i.e.*, σ*-additive*) *on* $\mathcal{R}$ *if and only if* μ *is continuous from below.*

Proof.

1) Let μ be a premeasure on $\mathcal{R}$ and $\{A_n\}$ a monotone nondecreasing sequence in $\mathcal{R}$ such that $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$. If $\mu(A_n) = \infty$ for some n, by monotonicity, $\mu(A_k) = \infty$ for all $k \geq n$ and $\mu(A) = \infty$. Then continuity from below follows immediately. Assume that $\mu(A_n) < \infty$ for all n. Denote

$$B_n = A_n\backslash A_{n-1}\,,\ \ n = 1, 2, \ldots,\ A_0 = \emptyset\,.$$

Then

$$A_n = \textstyle\sum_{k=1}^{n} B_k$$

and $\{B_n\}$ forms a pairwise disjoint sequence from $\mathcal{R}$ with

$$\textstyle\sum_{n=1}^{\infty} B_n = A.$$

Because $\mu(A_n) < \infty$ (by the above assumption), $\mu(B_n)$ is well defined and, therefore, due to σ-additivity of μ (as a premeasure),

$$\lim_{n\to\infty}\mu(A_n) = \lim_{n\to\infty}\sum_{k=1}^{n}\mu(B_k) = \sum_{k=1}^{\infty}\mu(B_k)$$

$$= \mu\left(\sum_{k=1}^{\infty}B_k\right) = \mu(A),$$

which yields continuity from below.

2) Now let μ be a content on $\mathcal{R}$, which is continuous from below, that is, for every monotone increasing sequence of sets in $\mathcal{R}$, $\{A_n\}\uparrow A := \bigcup_{n=1}^{\infty} A_n$ such that $A \in \mathcal{R}$, it holds that

$$\lim_{n\to\infty}\mu(A_n) = \mu(A).$$

Let $\{C_n\}$ be a sequence of pairwise disjoint sets. By setting

$$B_n = \sum_{k=1}^{n} C_k,$$

we get $\{B_n\}\uparrow \subseteq \mathcal{R}$, with $\bigcup_{n=1}^{\infty} B_n = B$ and hence $\lim_{n\to\infty}\mu(B_n) = \mu(B)$, which is equivalent to

$$\lim_{n\to\infty}\mu\left(\sum_{k=1}^{n}C_k\right) = \lim_{n\to\infty}\sum_{k=1}^{n}\mu(C_k) = \mu\left(\sum_{n=1}^{\infty}C_n\right).$$

This is the desired σ-additivity of μ. □

1.7 Theorem.

(i) *If μ is a premeasure on a ring $\mathcal{R}$ and if $\{A_n\}\downarrow \subseteq \mathcal{R}$ such that $\mu(A_1) < \infty$ and $A := \bigcap_{n=1}^{\infty} A_n \in \mathcal{R}$, then μ is Ø-continuous* (*or continuous from above*) *on sequence $\{A_n\}$. In particular, if μ is a finite premeasure on ring $\mathcal{R}$ then μ is Ø-continuous* (*or continuous from above*) *on $\mathcal{R}$.*

(ii) *If μ is a finite content on a ring $\mathcal{R}$ then Ø-continuity implies that μ is a premeasure.*

Proof.

(i). Because $\{A_n\}$ is monotone nonincreasing, $\{A_1 \backslash A_n\}\uparrow \subseteq \mathcal{R}$. It can easily be shown that $\bigcup_{n=1}^{\infty}(A_1 \backslash A_n) = A_1 \backslash \left(\bigcap_{n=1}^{\infty} A_n\right)$. Now we apply Lemma 1.6 to the sequence $\{A_1 \backslash A_n\}$ to arrive at

$$\lim_{n\to\infty} \mu(A_1 \backslash A_n) = \mu\left(\bigcup_{n=1}^{\infty}(A_1 \backslash A_n)\right) = \mu(A_1 \backslash A). \tag{1.7}$$

In as much as μ is finite on $\{A_n\}$, by Proposition 1.5(i), equation (1.7) reduces to

$$\mu(A_1) - \lim_{n\to\infty} \mu(A_n) = \mu(A_1) - \mu(A)$$

and thereby yields assertion (i).

(ii) We show that Ø-continuity implies Lemma 1.6 and that Lemma 1.6 in turn yields (ii). Let $\{A_n\}\uparrow \subseteq \mathcal{R}$ such that $\{A_n\}\uparrow A \in \mathcal{R}$. Then $\{A \backslash A_n\}\downarrow$ Ø and Ø-continuity yield

$$\lim_{n\to\infty} \mu(A \backslash A_n) = \mu(\text{Ø}) = 0.$$

Because the content μ is assumed to be finite on $\mathcal{R}$, by Proposition 1.5 (i), the last expression leads to

$$\lim_{n\to\infty} \mu(A_n) = \mu(A).$$

Applying Lemma 1.6 (2) we have that μ is a premeasure on $\mathcal{R}$. □

1.8 Examples.

(i) We show that on some measure spaces, while $\{A_n\}\downarrow$ Ø, $\mu(A_n) \not\to 0$. Let (Ω, Σ, μ) be such that $\Omega = \mathbb{N}$, $\Sigma = \mathcal{P}(\Omega)$, $\mu(\{n\}) = \frac{1}{n}$. Let $A_n = \{n, n+1, \dots\}$. Clearly, $\{A_n\}$ is a monotone decreasing vanishing sequence. However, $\mu(A_n) = \infty$ for each n and thus, $\mu(A_n) \to \infty$. So, Ø-continuity does not apply, because we violated the condition $\mu(A_1) < \infty$ of Theorem 1.7, which, as we see, is essential.

(ii) Consider in Ω the σ-algebra $\Sigma = \{\Omega, \emptyset, A, A^c\}$ and define

$$\mu : \Sigma \to \{1, 0, p, 1-p\},$$

where $p \in (0,1)$. Then μ is a (probability) measure on Σ, called a *Bernoulli measure*. Notice that for the traditional Bernoulli measure $\Omega = \{0,1\}$ and $A = \{1\}$.

(iii) Consider the following atomic measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$:

$$\mu = \sum_{k=0}^{n} a_k \varepsilon_k \,,$$

where

$$a_k = \binom{n}{k} p^k (1-p)^{n-k},\ p \in (0,1),\ k = 0, \ldots, n.$$

Clearly, μ is a probability measure. It is called the *binomial measure* (with parameters (n,p)) and it is denoted by $\beta_{n,p}$.

(iv) Consider another atomic measure on $\mathcal{B}(\mathbb{R})$:

$$\mu = \textstyle\sum_{n=0}^{\infty} a_n \varepsilon_n \,,\ \text{where } a_n = e^{-\alpha}\, \frac{\alpha^n}{n!},\ \alpha > 0,\ n = 0, 1, \ldots.$$

μ is also a probability measure, called the *Poisson measure* (with parameter α), in notation, π_α. □

1.9 Notation. Denote by $\mathfrak{M}(\Omega, \Sigma)$ the set of all measures on a measurable space (Ω, Σ). According to Problem 1.4 [Example 1.2 (iv)], any positive linear combination $\alpha_1\mu_1 + \alpha_2\mu_2$ of measures (that is, such that $\alpha_1, \alpha_2 \in \mathbb{R}_+$) is a measure. With $\alpha = 0$, we have $\alpha\mu = \theta$, being zero measure if we agree $0 \cdot \infty = 0$, which is commonly agreed upon in measure theory. So, with $\theta \in \mathfrak{M}(\Omega, \Sigma)$, $\mathfrak{M}(\Omega, \Sigma)$ is a semilinear space over the semifield $\mathbb{R}_+$, in notation $(\mathfrak{M}(\Omega, \Sigma),\ \mathbb{R}_+,\ +\,,\ \cdot\,)$. We show that this space of measures is also a lattice, under a particular lattice operation. □

1.10 Lemma. *Let $\mu_1, \mu_2 \in \mathfrak{M}(\Omega, \Sigma)$. Then the set function*

$$\rho(A) = \sup\{\mu_1(E) + \mu_2(A \backslash E) : E \in \Sigma \cap A\}$$

is a measure on Σ.

Proof. Obviously, ρ is a nonnegative set function defined on Σ that equals zero at the empty set.

Next we show that ρ is monotone. Let $E \subseteq A \subseteq B$. Then, we have

$$\mu_1(E) + \mu_2(A\backslash E) \leq \mu_1(E) + \mu_2(B\backslash E).$$

Because the inequality holds true for all $E \subseteq A$ and $E \subseteq B$, $\rho(A) \leq \rho(B)$.

We need to prove that ρ is σ-additive. Let $\{A_n\} \subseteq \Sigma$ be a sequence of disjoint sets with $A := \sum_{n=1}^{\infty} A_n$. We will first show that ρ is σ-*subadditive* on $\{A_n\}$, that is,

$$\rho(A) \leq \sum_{n=1}^{\infty} \rho(A_n). \tag{1.10}$$

Let E be a measurable subset of A. Then,

$$\begin{aligned} &\mu_1(E) + \mu_2(A\backslash E) \\ &= \mu_1\left(\sum_{n=1}^{\infty}(A_n \cap E)\right) + \mu_2\left(\sum_{n=1}^{\infty}(A_n \cap E^c)\right) \\ &= \sum_{n=1}^{\infty} \mu_1(A_n \cap E) + \mu_2(A_n\backslash(A_n \cap E)) \\ &\leq \sum_{n=1}^{\infty} \rho(A_n) \end{aligned}$$

which yields (1.10).

It remains to show the inverse inequality. To avoid a trivial case, we assume that $\rho(A) < \infty$. Then, by monotonicity, $\rho(A_n) \leq \rho(A) < \infty$. From the definition of ρ, given A_n, and $\varepsilon_n > 0$, there is a subset $E_n \subseteq A_n$ such that

$$\rho(A_n) < \mu_1(E_n) + \mu_2(A_n\backslash E_n) + \varepsilon_n,\ n = 1, 2, \ldots .$$

We choose ε_n's such that $\sum_{n=1}^{\infty} \varepsilon_n \leq \varepsilon$, which we choose beforehand. Denote $E = \sum_{n=1}^{\infty} E_n$. Obviously, E_n's are disjoint and $E \subseteq A$ and

$$A\backslash E = \sum_{n=1}^{\infty}(A_n\backslash \sum_{k=1}^{\infty} E_k) = \sum_{n=1}^{\infty}(A_n\backslash E_n),$$

because the rest of E_k's are disjoint with A_n. Therefore, with

$$\mu_1(E) + \mu_2(A\backslash E) = \sum_{n=1}^{\infty}\mu_1(E_n) + \sum_{n=1}^{\infty}\mu_2(A_n\backslash E_n)$$

we have:

$$\sum_{n=1}^{\infty}\rho(A_n) \le \sum_{n=1}^{\infty}\mu_1(E_n) + \sum_{n=1}^{\infty}\mu_2(A_n\backslash E_n) + \sum_{n=1}^{\infty}\varepsilon_n$$

$$\le \mu_1(E) + \mu_2(A\backslash E) + \varepsilon \le \rho(A) + \varepsilon.$$

This inequality holds for each $\varepsilon > 0$. Therefore, we arrive at the targeted inverse inequality and we are done with the lemma. □

1.11 Lemma. *Let $\mu_1, \mu_2 \in \mathfrak{M}(\Omega, \Sigma)$. Then the set function*

$$\nu(A) = \inf\{\mu_1(E) + \mu_2(A\backslash E) : E \in \Sigma \cap A\}$$

is a measure on Σ.

The proof of this lemma is similar to that of Lemma 1.12 and is left for the reader (Problem 1.16). □

1.12 Notation. We introduce on $\mathfrak{M}(\Omega, \Sigma)$ the partial order $\preceq$ such that $\mu \preceq \nu$ if and only if $\mu(A) \le \nu(A)$ for all $A \in \Sigma$. □

The following theorem states that $(\mathfrak{M}(\Omega, \Sigma), \preceq)$ is a lattice.

1.13 Theorem. *Let ρ and ν be defined as in Lemmas* 1.10 *and* 1.11. *Then, $\rho = \mu_1 \vee \mu_2$ and $\nu = \mu_1 \wedge \mu_2$, and consequently, the partially ordered set $(\mathfrak{M}(\Omega, \Sigma), \preceq)$ is a lattice.*

Proof. (i) Let $\mathcal{U}(\{\mu_1, \mu_2\}) \subseteq \mathfrak{M}(\Omega, \Sigma)$ denote the set of all upper bounds of the two-element set $\{\mu_1, \mu_2\}$. We show that $\rho \in \mathcal{U}(\{\mu_1, \mu_2\})$. In other words, we show that for each $A \in \Sigma$, $\rho(A) \ge \mu_1(A)$ and $\rho(A) \ge \mu_2(A)$, in notation, $\rho \succeq \mu_1$ and $\rho \succeq \mu_2$. Indeed, for any measurable subset $E \subseteq A$,

$$\rho(A) \ge \mu_1(E) + \mu_2(A\backslash E).$$

Choosing $E = A$ or $E = \emptyset$, we arrive at $\rho \succeq \{\mu_1, \mu_2\}$ and thus $\rho \in \mathcal{U}(\{\mu_1, \mu_2\})$. Similarly, it can be shown that the measure ν from Lemma 1.11 belongs to the set $\mathcal{L}(\{\mu_1, \mu_2\})$ of all lower bounds of $\{\mu_1, \mu_2\}$, in notation, $\nu \preceq \{\mu_1, \mu_2\}$.

(ii) Let $u \in \mathcal{U}(\{\mu_1, \mu_2\})$ be another upper bound of the set $\{\mu_1, \mu_2\}$. We show that $\rho \preceq u$, which validates that ρ is the smallest upper bound of $\{\mu_1, \mu_2\}$. Let $E \subseteq A$ be two measurable sets. Then, we have

$$\mu_1(E) + \mu_2(A\backslash E) \leq u(E) + u(A\backslash E) = u(A)$$

implying that $\rho \preceq u$ and that $\rho = \mu_1 \vee \mu_2$.

Now, let $\lambda \in \mathcal{L}(\{\mu_1, \mu_2\})$. Similarly,

$$\lambda(A) = \lambda(E) + \lambda(A\backslash E) \leq \mu_1(E) + \mu_2(A\backslash E) \tag{1.13}$$

holding for all $E \subseteq A$. Therefore, the inequality will sustain for the infimum on the right of (1.13). Consequently, $\lambda \preceq \nu$ and thus ν is the largest lower bound of $\{\mu_1, \mu_2\}$. □

1.14 Remark. In light of Definition 7.9(ii), Chapter 1, $(\mathfrak{M}(\Omega, \Sigma), \mathbb{R}_+, +, \cdot, \preceq)$ is a semivector lattice. □

PROBLEMS

1.1 Let $\mathcal{S}$ be the semi-ring of half open intervals on the real line and λ^0 be the Lebesgue elementary content. Take $A = (0, 1]$, $B = (1, 2]$, and $C = (3, 4]$. Although $\lambda^0(A + B) = \lambda^0(A) + \lambda^0(B)$, we cannot state that $\lambda^0(A + C) = \lambda^0(A) + \lambda^0(C)$, because $A + C$ is not an interval, and therefore, the left-hand side of the last equation is not defined. Hence, λ^0 is not additive on $\mathcal{S}$. True of false?

1.2 Prove Proposition 1.5(i, ii). *Every content μ on a ring $\mathcal{R}$ has the following properties*:

(i) $A \subseteq B \Rightarrow \mu(B\backslash A) = \mu(B) - \mu(A)$ [provided that $\mu(A) < \infty$].

(ii) $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

1.3 Show that for a content on a ring, the notions of continuity from below and Ø-continuity are equivalent.

1.4 Prove that μ is a measure on Σ in Example 1.2(v).

1.5 Let (Ω, Σ, μ) be a measure space and let $\{A_n : n = 1, 2, \dots\} \subseteq \Sigma$ with $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Show that $\mu\big(\varlimsup_{n \to \infty} A_n\big) = 0$ [*Hint*: Apply continuity from above of measure μ to the monotone nonincreasing sequence $\left\{ \bigcup_{k=n}^{\infty} A_k : n = 1, 2, \dots \right\}$.]

1.6 A subset $\mathcal{R}_I$ of a ring $\mathcal{R}$ in Ω is called an *ideal in* $\mathcal{R}$ if it has the following properties:

$a)$ $\emptyset \in \mathcal{R}_I$.

$b)$ $B \in \mathcal{R}_I,\ A \subseteq B,\ A \in \mathcal{R} \ \Rightarrow\ A \in \mathcal{R}_I$.

$c)$ $A,\ B \in \mathcal{R}_I \ \Rightarrow\ A \cup B \in \mathcal{R}_I$.

Let μ be a content on a ring $\mathcal{R}$. Define $\mathcal{R}_\mu = \{R \in \mathcal{R} : \mu(R) = 0\}$. Show that $\mathcal{R}_\mu$ is an ideal in $\mathcal{R}$.

1.7 A subset $\mathcal{R}_{\sigma\text{-}I}$ of a ring $\mathcal{R}$ is called a *σ-ideal in* $\mathcal{R}$ if, in addition to properties a-$c)$ of Problem 1.6, it is closed with respect to the formation of at most countable unions. Let μ be a premeasure on $\mathcal{R}$. With the same definition of $\mathcal{R}_\mu$ as in Problem 1.6, prove that $\mathcal{R}_\mu$ is a σ-ideal in $\mathcal{R}$.

1.8 Let $\mathcal{R}_{\sigma\text{-}I}$ be a σ-ideal in a ring $\mathcal{R}$. Show that there exists a premeasure μ on $\mathcal{R}$ such that $\mathcal{R}_{\sigma\text{-}I} = \mathcal{R}_\mu$, defined in Problems 1.7 and 1.8.

1.9 Let μ be a finite content on a ring $\mathcal{R}$. Show that $d(A, B) := \mu(A \triangle B)$, for all $A, B \in \mathcal{R}$, is a pseudo-metric on $\mathcal{R}$ [that is, that d possesses all properties of a metric except for $d(A, B) = 0$ yields that $A = B$].

1.10 Let (Ω, Σ, μ) be a measure space and $\{A_n\} \subseteq \Sigma$ be a sequence of sets with $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) < \infty$ and $\inf\{\mu(A_n)$: $n = 1, 2, \dots\} = a \geq 0$. Show that $\mu\big(\overline{\lim}_{n \to \infty} A_n\big) \geq a$.

1.11 Let (Ω, Σ, μ) be a measure space and for a number $0 < a \leq \infty$, define $\mathcal{G}_a = \{G \in \Sigma : \mu(G) < a\}$ and

$$\Sigma_a = \{Q \subseteq \Omega : Q \cap G \in \Sigma,\ \forall G \in \mathcal{G}_a\}.$$

Show that Σ_a is a σ-algebra.

1.12 Under the condition of Problem 1.11, let $a = \infty$ and μ be a finite measure. Show that $\Sigma_\infty = \Sigma$.

1.13 Let (Ω, Σ, μ) be a measure space, $\mathcal{G}_\infty = \{G \in \Sigma : \mu(G) < \infty\}$, and $\Sigma_\infty = \{Q \subseteq \Omega : Q \cap G \in \Sigma, \forall G \in \mathcal{G}_\infty\}$. Define the set function μ_∞ on Σ_∞ as

$$\mu_\infty(Q) = \begin{cases} \mu(Q), & Q \in \Sigma \\ \infty, & Q \notin \Sigma. \end{cases}$$

(Notice that $\Sigma \subseteq \Sigma_\infty$.) Show that μ_∞ is a measure on Σ_∞.

1.14 Argue that for any probability space (Ω, Σ, μ), the axiom $\mu(\emptyset) = 0$ is redundant. Is it also true for any measure?

1.15 Show that for each two real numbers $a < b$, $f^*((a, b])$ in Example 1.2 (iii) is a Borel set.

1.16 Prove Lemma 1.11. *Let* $\mu_1, \mu_2 \in \mathfrak{M}(\Omega, \Sigma)$. *Then the set function*

$$\nu(A) = \inf\{\mu_1(E) + \mu_2(A \backslash E) : E \in \Sigma \cap A\}$$

is a measure on Σ.

1.17 Prove the following statement. *Let* $\{\mu_n\} \subseteq \mathfrak{M}(\Omega, \Sigma)$ *be a monotone nondecreasing sequence of measures in the sense of* $\preceq$. *Then the pointwise limit of the sequence* $\mu := $

$\lim_{n\to\infty}\mu_n$ (*that is,* $\mu(A) = \lim_{n\to\infty}\mu_n(A), A \in \Sigma$), *is a measure on* Σ *and* $\mu = \sup\{\mu_n\}$.

1.18 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and let $\mathcal{G} \subseteq \Sigma$ be a family of events. $\mathcal{G}$ is called *independent* if for any finite tuple of events $G_1, \ldots, G_m$, $\mathbb{P}\left(\bigcap_{i=1}^{m} G_i\right) = \prod_{i=1}^{m} \mathbb{P}(G_i)$. Let $\mathcal{D}(\mathcal{G})$ be the Dynkin system induced by $\mathcal{G}$. Show that $\mathcal{D}(\mathcal{G})$ is also independent.

1.19 Under the condition of Problem 1.18, let $\mathcal{G}$ be a semi-ring. Then show that $\Sigma(\mathcal{G})$, the σ-algebra induced by $\mathcal{G}$, is independent.

1.20 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and let $\{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ be a set of n families of events. This set is said to be *independent* if for any choice of events $G_1 \in \mathcal{G}_1, \ldots, G_n \in \mathcal{G}_n$, the tuple $\{G_1, \ldots, G_n\}$ is independent. If the set $\{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ is independent, show that the set $\{\mathcal{D}(\mathcal{G}_1), \ldots, \mathcal{D}(\mathcal{G}_n)\}$ (of Dynkin systems induced by $\mathcal{G}_i$'s) is also independent.

1.21 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and let $\{\mathcal{G}_i : i \in I\}$ be an indexed system of families of events. This system is said to be *independent* if for whatever choice of n indices, $i_1, \ldots, i_n \in I$ and then a choice of n events $G_1 \in \mathcal{G}_{i_1}, \ldots, G_n \in \mathcal{G}_{i_n}$, the tuple $\{G_1, \ldots, G_n\}$ is independent. If the set $\{\mathcal{G}_i : i \in I\}$ is independent, show that the set $\{\mathcal{D}(\mathcal{G}_i) : i \in I\}$ (of Dynkin systems induced by $\mathcal{G}_i$'s) is also independent.

1.22 Under the condition of Problem 1.20, suppose that $\mathcal{G}_i$'s are semi-rings. Show that the set $\{\Sigma(\mathcal{G}_1), \ldots, \Sigma(\mathcal{G}_n)\}$ is independent.

NEW TERMS:

Peano, Giuseppe, 261
Jordan, Camille, 261
Borel, Émile, 262
Kolmogorov, Andrey, 262
Lebesgue, Henri, 262
Vitali, Giuseppe, 262
Radon, Johann, 262
set function, 263
finitely additive set function, 263
additive set function, 263
σ-additive set function, 263
continuity from below on a σ-algebra, 263
continuity from below on a sequence of sets, 263
continuity from above on a sequence of sets, 264
continuity from above on a σ-algebra, 264
continuity from above at the empty set, 264
Ø-continuity, 264
finite set function, 264
σ-finite set function on a system of sets, 264
σ-finite set function on a sequence of sets, 264
elementary content, 264
content, 264
premeasure, 264
measure, 264
probability measure, 264
measure space, 264
probability space, 264
point mass (Dirac measure), 265
Dirac measure (point mass), 265
Lebesgue elementary content, 265
distribution function, 265
extended distribution function, 265
Lebesgue - Stieltjes elementary content, 266
atomic (discrete) measure, 267
discrete (atomic) measure, 267

atomic probability measure, 267
counting measure, 268
counting measure space, 268
monotonicity of measures, 268
μ-minimal decomposition of a set, 270
σ-subadditivity, 270
finite subadditivity, 270
continuity from below of a premeasure, criterion of, 271
continuity from above of a premeasure, criterion of, 272
Bernoulli measure, 274
binomial measure, 274
Poisson measure, 274
set of measures as a lattice, 276
semivector lattice of measures, 277
ideal, 278
σ-ideal, 278
pseudo-metric on a ring, 278
independent families of events, 280
independent Dynkin systems of events, 280
independent σ-algebras of events, 280

2. EXTENSION OF SET FUNCTIONS TO A MEASURE

We begin this section with the introduction of a set function that is not exactly a measure, as it is not even additive, but which is at the heart of the formation of measures extended from some more primitive set functions. A prominent example of such a construction yields the Lebesgue measure. It is firstly defined on "rectangles" and rectangular figures. Then the measurement of a more general set is accomplished by an approximation of the set by rectangles that cover the set. The latter leads to the notion of an "outer measure," which was initially proposed by Lebesgue at the turn of the last century and later on refined by Carathéodory. Carathéodory's approach is essentially preserved in the contemporary constructions.

In the four steps below we outline the Carathéodory extension procedure.

1. An Outer Measure. We first introduce an *outer measure* μ^* on $\mathcal{P}(\Omega)$ that is a monotone set function ($A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$) and σ-*subadditive*:

$$\{Q_n\} \subseteq \mathcal{P}(\Omega) \Rightarrow \mu^*\left(\bigcup_{n=1}^{\infty} Q_n\right) \leq \sum_{n=1}^{\infty} \mu^*(Q_n).$$

Secondly, we identify the family of μ^*-*measurable sets* $\Sigma^* \subseteq \mathcal{P}(\Omega)$ that exhibit some sort of additivity property. Namely, for an $M \in \Sigma^*$ (because M separates any $Q \subseteq \Omega:\ Q = Q \cap M + Q \cap M^c$) we have $\mu^*(Q) = \mu^*(Q \cap M) + \mu^*(Q \cap M^c)$.

It turns out that Σ^* is a σ-algebra in Ω and $\mu_0^* = \mathrm{Res}_{\Sigma^*}\mu^*$ is a measure on Σ^*. Moreover, this measure is *complete*. The latter is often a valuable property on its own (utilized in many applications); namely, if $\mu_0^*(N) = 0$ for some Σ^*-measurable set (null set), then $\mathcal{P}(N) \subseteq \Sigma^*$, which in general, need not be measurable.

2. The Construction of an Outer Measure by a Formatter. Let $\mathcal{G}$ be a fixed family of subsets of Ω identified as *rectangles* (by the analogue with rectangles in $\mathbb{R}^n$ and their role in measuring sets) and γ be a set function on $\mathcal{G}$. For any subset $Q \subseteq \Omega$, we define the set $\mathscr{C}_Q = \mathscr{C}_Q(\mathcal{G})$ of all at most countable covers of Q by rectangles. Then, it turns out that the set function μ^* on $\mathcal{P}(\Omega)$ defined as

$$\mu^*(Q) = \begin{cases} \inf\{\sum_{n=1}^{\infty}\gamma(G_n)\colon \{G_n\} \in \mathscr{C}_Q\}, & \mathscr{C}_Q \neq \emptyset \\ \infty, & \mathscr{C}_Q = \emptyset \end{cases}$$

is an outer measure. We call the pair $(\mathcal{G}, \gamma)$ a *formatter* of the outer measure μ^*.

3. The Extension of Formatter $(\mathcal{G}, \gamma)$ to $(\Sigma(\mathcal{G}), \mu_0^*)$. A reasonable way to enlarge a formatter (from a measure-theoretical point of view) is to attain the more populous $\Sigma(\mathcal{G})$ and then have some measure on $\Sigma(\mathcal{G})$, say μ, extended from γ (that is, such that $\text{Res}_{\mathcal{G}}\mu = \gamma$). A desirable extension must clearly be unique. In light of Steps 1 and 2, accomplishing it in a plausible way suggests the formation of the outer measure μ^* followed by the reduction of $\mathcal{P}(\Omega)$ to Σ^* on which μ^* is a measure.

But does this process leads to $\Sigma(\mathcal{G})$ (which seems to be a smaller σ-algebra containing $\mathcal{G}$)? Unfortunately Σ^* is unrelated to $\Sigma(\mathcal{G})$, because $\mathcal{G} \nsubseteq \Sigma^*$, and furthermore, $\mu^* \neq \gamma$ on $\mathcal{G}$, unless we make some assumptions about $\mathcal{G}$ and γ. If such conditions exist, we call formatter $(\mathcal{G}, \gamma)$ *extendible* (through the outer measure it produces). If the formatter is extendible, then (Σ^*, μ_0^*) is called the *complete Carathéodory extension* of $(\mathcal{G}, \gamma)$ and its restriction to $(\Sigma(\mathcal{G}), \mu_0^*)$ is called a *Carathéodory extension*.

4. The "Extendibility" Conditions for Formatter $(\mathcal{G}, \gamma)$. If the formatter $(\mathcal{G}, \gamma)$ carries a semi-ring and elementary content, respectively, then $\mathcal{G} \subseteq \Sigma^*$. If γ is σ-additive on $\mathcal{G}$, then (Σ^*, μ_0^*) becomes the complete Carathéodory extension of $(\mathcal{G}, \gamma)$. Consequently, the restriction from (Σ^*, μ_0^*) to $(\Sigma(\mathcal{G}), \mu_0^*)$ is a Carathéodory extension of the formatter. In other words, there is a measure on $\Sigma(\mathcal{G})$ that coincides with γ on $\mathcal{G}$. Finally, if not only is γ σ-additive but also σ-finite, the Carathéodory extension is unique. □

We now turn to the details of the above procedure followed by a discussion of various formatters.

2.1 Definition. Let Ω be a nonempty set and μ^* be a set function defined on $\mathcal{P}(\Omega)$. μ^* is called an *outer measure* if:

$a)\ \mu^*(\emptyset) = 0$ (a reminder). (2.1a)

$b)\ A \subseteq B \ \Rightarrow\ \mu^*(A) \le \mu^*(B)$ (*monotonicity*). (2.1b)

$$c)\ \{Q_n\} \subseteq \mathcal{P}(\Omega) \ \Rightarrow \mu^*\left(\bigcup_{n=1}^{\infty} Q_n\right) \le \sum_{n=1}^{\infty} \mu^*(Q_n)\ (\sigma\text{-}\textit{subadditivity}). \tag{2.1c}$$

□

2.2 Definition. Let μ^* be an outer measure on $\mathcal{P}(\Omega)$. A subset $M \subseteq \Omega$ is said to be μ^*-*measurable*, if for any $Q \subseteq \Omega$,

$$\mu^*(Q) = \mu^*(Q \cap M) + \mu^*(Q \cap M^c). \tag{2.2}$$

We also say that M μ^*-*separates* Q. □

The following is what essentially constitutes the widely referred to Carathéodory extension theorem that we outlined above. For convenience, we break it up into several theorems. The idea of an outer measure, along with the extension, belongs to the German mathematician (of the Greek origin) Constantin Carathéodory that appeared in his 1914 paper, "Über das lineare Mass von Punktmengen - eine Verallgemeinerung des Längebegriffes" (in *Göttingen Nachrichten*) and in his widely referred-to 1918 book, *Vorlesungen über Reellen Funktionen* (Teubner, Leipzig). We start with the following.

2.3 Theorem. *The collection Σ^* of all μ^*-measurable subsets is a σ-algebra in Ω. The restriction of μ^* from $\mathcal{P}(\Omega)$ to Σ^*, in notation μ_0^*, is a measure*.

Proof. Because throughout the proof of this theorem we largely use equation (2.2) or prove its validity, we first notice that, due to σ-subadditivity of μ^*, as an outer measure, the inequality

$$\mu^*(Q) \le \mu^*(Q \cap M) + \mu^*(Q \cap M^c) \tag{2.3}$$

holds true for all subsets, Q and M, of Ω.

Our proof includes the following steps.

a) Ω is obviously an element of Σ^*, as it satisfies (2.2). If $M \in \Sigma^*$, then $M^c \in \Sigma^*$, by their symmetry in (2.2).

b) We show that Σ^* is closed with respect to the formation of finite unions, that is, we show that with $A, B \in \Sigma^*$, $A \cup B \in \Sigma^*$.

Because $B \in \Sigma^*$, it follows that for each $Q' \in \mathcal{P}(\Omega)$,

$$\mu^*(Q') = \mu^*(Q' \cap B) + \mu^*(Q' \cap B^c). \tag{2.3a}$$

Specifically, (2.3a) is valid for $Q' = Q \cap A$ and $Q' = Q \cap A^c$, $Q \in \mathcal{P}(\Omega)$. Hence,

$$\mu^*(Q \cap A) = \mu^*(Q \cap A \cap B) + \mu^*(Q \cap A \cap B^c)$$

and

$$\mu^*(Q \cap A^c) = \mu^*(Q \cap A^c \cap B) + \mu^*(Q \cap A^c \cap B^c).$$

Summing up the last two equations and taking into account that $A \in \Sigma^*$, we have

$$\begin{aligned}\mu^*(Q) &= \mu^*(Q \cap A) + \mu^*(Q \cap A^c)\\ &= \mu^*(Q \cap A \cap B) + \mu^*(Q \cap A \cap B^c)\\ &+ \mu^*(Q \cap A^c \cap B) + \mu^*(Q \cap A^c \cap B^c),\end{aligned}$$

implying that

$$\begin{aligned}\mu^*(Q) &= \mu^*(Q \cap A \cap B) + \mu^*(Q \cap A \cap B^c)\\ &+ \mu^*(Q \cap A^c \cap B) + \mu^*(Q \cap (A \cup B)^c).\end{aligned} \tag{2.3b}$$

Now replacing Q in (2.3b) with $Q \cap (A \cup B)$ we also have

$$\begin{aligned}&\mu^*(Q \cap (A \cup B))\\ &= \mu^*(Q \cap (A \cup B) \cap A \cap B) + \mu^*(Q \cap (A \cup B) \cap A \cap B^c)\end{aligned}$$

$$+\mu^*(Q\cap\ (A\cup B)\cap B\cap A^c)+\mu^*(Q\cap(A\cup B)\cap A^c\cap B^c)$$

$$=\mu^*(Q\cap A\cap B)+\mu^*(Q\cap A\cap B^c)$$

$$+\mu^*(Q\cap B\cap A^c)+\mu^*(Q\cap\emptyset).$$

The latter reduces to

$$\mu^*(Q\cap(A\cup B))=\mu^*(Q\cap A\cap B)$$

$$+\mu^*(Q\cap A\cap B^c)+\mu^*(Q\cap B\cap A^c). \tag{2.3c}$$

Substituting (2.6) into (2.5) we get

$$\mu^*(Q)=\mu^*[Q\cap(A\cup B)]+\mu^*[Q\cap(A\cup B^c)]\,,$$

which shows that $A\cup B\in\Sigma^*$.

The above assertions a) and b) imply that Σ^* is an algebra in Ω.

c) Now we prove that Σ^* is a σ-algebra in Ω. Because Σ^*, as an algebra, is $\cap$-stable, it is sufficient to show that Σ^* is a Dynkin system. (See Problem 1.10 of Chapter 4.)

Let $\{A_n\}\subseteq\Sigma^*$ be a sequence of disjoint sets. Take $A_1, A_2\in\{A_n\}$. Substituting $A_1=A$ and $A_2=B$ into (2.3c), taking A and B in (2.3c) disjoint, and then noticing that $A\cap B^c=A$ and $B\cap A^c=B$, we arrive at

$$\mu^*[Q\cap(A+B)]=\mu^*(Q\cap A)\ +\mu^*(Q\cap B). \tag{2.3d}$$

If $A_1,\ldots,A_n$ is an n-tuple of mutually disjoint elements of Σ^*, then, by induction, from (2.3d),

$$\mu^*[Q\cap S_n]=\textstyle\sum_{k=1}^n\mu^*(Q\cap A_k), \tag{2.3e}$$

where $S_n=\sum_{k=1}^n A_k$. Denote $S=\sum_{n=1}^\infty A_n$. Because of $S_n\subseteq S$, $(Q\cap S^c)\subseteq(Q\cap S_n^c)$, and by monotonicity of μ^*,

$$\mu^*(Q \cap S^c) \leq \mu^*(Q \cap S_n^c). \tag{2.3f}$$

Because Σ^* is an algebra, it follows that $S_n \in \Sigma^*$ and hence it is μ^*-measurable, i.e., it μ^*-separates Q, which, combined with (2.3e) and (2.3f), yields

$$\begin{aligned}\mu^*(Q) &= \mu^*(Q \cap S_n) + \mu^*(Q \cap S_n^c) \\ &\geq \sum_{k=1}^{n} \mu^*(Q \cap A_k) + \mu^*(Q \cap S^c),\ n = 1, 2, \ldots.\end{aligned}$$

Therefore,

$$\mu^*(Q) \geq \textstyle\sum_{k=1}^{\infty} \mu^*(Q \cap A_k) + \mu^*(Q \cap S^c) \tag{2.3g}$$

that, by σ-subadditivity, gives

$$\begin{aligned}&\geq \mu^*(Q \cap \textstyle\sum_{k=1}^{\infty} A_k) + \mu^*(Q \cap S^c) \\ &= \mu^*(Q \cap S) + \mu^*(Q \cap S^c).\end{aligned} \tag{2.3h}$$

Inequalities (2.3) and (2.3g - 2.3h) lead to

$$\mu^*(Q) = \mu^*(Q \cap S) + \mu^*(Q \cap S^c)$$

concluding that $S = \sum_{n=1}^{\infty} A_n$ indeed μ^*-separates any $Q \subseteq \mathcal{P}(\Omega)$ and thus is an element of Σ^*. The latter supports the claim that Σ^* is a Dynkin system and, consequently, that Σ^* is a σ-algebra.

d) We show that μ_0^* is a measure on Σ^*. Substituting the set $S = \sum_{n=1}^{\infty} A_n$ for Q in (2.3g), we have

$$\mu_0^*(\textstyle\sum_{n=1}^{\infty} A_n) \geq \sum_{k=1}^{\infty} \mu_0^*(A_k),$$

which, due to σ-subadditivity of μ^*, leads to the strict equality and thereby, σ-additivity of μ_0^*. Therefore, we have proved that $\mathrm{Res}_{\Sigma^*}\mu^*$, denoted by μ_0^*, is a measure.

The proof is, therefore, completed. □

2.4 Examples.

(i) Let

$$\Omega = \{a, b, c\},\ A = \{a\},\ A^c = \{b, c\},$$

$$P = \{b\},\ Q = \{c\},\ R = \{a, b\},\ S = \{a, c\}.$$

Define the following set function μ^* on $\mathcal{P}(\Omega)$.

$$\mu^*(\emptyset) = 0,\ \mu^*(\Omega) = 4,\ \mu^*(A) = 1,$$

$$\mu^*(A^c) = \mu^*(R) = \mu^*(S) = 3,\ \mu^*(P) = \mu^*(Q) = 2.$$

One can easily verify that μ^* is an outer measure on

$$\mathcal{P}(\Omega) = \{\emptyset, \Omega, A, A^c, P, Q, R, S\},$$

as it satisfies axioms (2.1a - 2.1c), but μ^* is not a measure, because it is not additive. We can see that only the sets $\emptyset$, Ω, A, and A^c μ^*-separate all subsets of Ω and, consequently, $\{\emptyset, \Omega, A, A^c\}$ is the σ-algebra Σ^*. Clearly, μ_0^*, as the restriction of μ^* on Σ^*, is a measure.

(ii) Let Ω be an infinite set. Define the set function γ on $\mathcal{P}(\Omega)$ by $\gamma(Q) = 0$ if Q is a finite set and $\gamma(Q) = 1$ if Q is infinite. Let $Q = \{\{\omega_n\},\ n = 1, 2, \dots\}$ be a sequence of all different singletons. Then,

$$\sum_{n=1}^{\infty} \gamma(\{\omega_n\}) = 0,$$

and $\gamma(Q) = 1$. Thus, γ is not σ-subadditive and not an outer measure. □

Recall that a *restriction* of a function $[X, Y, f]$ is a function $[X_0, Y_0, f_0]$ defined on contracted domain $X_0 \subseteq X$ with $f = f_0$ on X_0 and $Y_0 \subseteq Y$. (In notation, $f_0 = \mathrm{Res}_{X_0} f$.) From Theorem 2.3, we learned that the set function $[\Sigma^*, [0, \infty], \mu_0^*]$ is a restriction of outer measure $[\mathcal{P}(\Omega), [0, \infty], \mu^*]$.

If $\overline{X}$ and $\overline{Y}$ are supersets of X and Y, respectively, a function $[\overline{X}, \overline{Y}, \overline{f}]$ is called an *extension* of f (*from* X *to* $\overline{X}$), if $[X, Y, f]$ is the restriction of $\overline{f}$ to X. (In notation, $\overline{f} = \mathrm{Ext}_{\overline{X}} f$.) We will apply this notion to extend a set function γ defined on a collection $\mathcal{G}$ of subsets of Ω to a set function $\overline{\gamma}$ on an expanded family $\mathcal{E}(\mathcal{G})$ of subsets of Ω. For instance, in Example $1.2(ii)$ we defined the Lebesgue elementary content λ^0 on the semi-ring $\mathcal{S}$ of half-open intervals in $\mathbb{R}^n$. We can extend the Lebesgue elementary content λ^0 to a (unique) content λ_c on $\mathcal{R}(\mathcal{S})$ (see Problem 2.2), which turns out to be a premeasure on $\mathcal{R}$ (verified in Theorem 3.1). The chief objective in this section is to construct an extension of a set function, such as premeasure, given on a ring, to a measure on the smallest σ-algebra generated by this ring. Other matters, such as "completion" of a measure, will also be a focus of our discussions.

2.5 Definitions.

(i) Let (Ω, Σ, μ) be a measure space. A measurable set $N \in \Sigma$ is called a *μ-null set* (or just *null set*) if $\mu(N) = 0$. We denote the set of all μ-null sets by $\mathcal{N}_\mu$. A set E is called *μ-negligible* (or just *negligible*), if there is a measurable null superset of E. (Note that a negligible set need not be measurable.) The measure space is called *complete*, if for each null set $N \in \mathcal{N}_\mu$, $\mathcal{P}(N) \subseteq \Sigma$, that is, if all negligible sets are measurable.

(ii) Consider a measure space (Ω, Σ, μ). Let $\overline{\Sigma}$ be the collection of all sets of type $A \cup M$ where $A \in \Sigma$ and M is any negligible set. According to Problem 2.8, $\overline{\Sigma}$ is a σ-algebra. We extend μ to $\overline{\mu}$ on $\overline{\Sigma}$ by setting

$$\overline{\mu}(A \cup M) = \overline{\mu}(A) = \mu(A).$$

According to Problem 2.8, the extension $(\overline{\Sigma}, \overline{\mu})$ of (Σ, μ) (or just $\overline{\mu}$) is complete. Measure $\overline{\mu}$ is then called the *completion of measure* μ and the measure space $(\Omega, \overline{\Sigma}, \overline{\mu})$ is called the *completion of measure space* (Ω, Σ, μ). □

2.6 Example. Let $\Omega = \mathbb{R}$, $\Sigma = \{A \in \mathcal{P}(\mathbb{R}) : \text{either } A \text{ or } A^c \preceq \mathbb{N}\}$, which is a σ-algebra on Ω [see Example 1.2 (vii),

Chapter 4], and let ε_1 be the point mass. Both sets $A = \{\frac{1}{n},\ n = 1, 2, \dots\}$ and A^c are elements of Σ and $\varepsilon_1(A^c) = 0$. Obviously, $E = [2, \infty)$, as a subset of A^c, is negligible, but not measurable. Therefore, the measure space $(\mathbb{R}, \Sigma, \varepsilon_1)$ is not complete. (See a more general case in Problem 2.14.) □

The proposition below is a paradigm of a complete measure space.

2.7 Proposition. *The restriction μ_0^* of an outer measure μ^* to the σ-algebra Σ^* of all μ^*-measurable subsets of Ω is complete and $(\Omega, \Sigma^*, \mu_0^*)$ is a complete measure space.*

Proof. Because μ^* is defined on whole $\mathcal{P}(\Omega)$, for any μ^*-negligible subset $N \subseteq \Omega$, due to (2.1b), $\mu^*(N) = 0$ and, therefore, it is sufficient to show that N is μ^*-measurable. Let $Q \subseteq \Omega$. Due to monotonicity of the outer measure, $\mu^*(Q \cap N) = 0$ and $\mu^*(Q \cap N^c) \leq \mu^*(Q)$ and this, along with (2.3), yields

$$\mu^*(Q) \leq \mu^*(Q \cap N) + \mu^*(Q \cap N^c) \leq \mu^*(Q)$$

and, hence, that $N \in \Sigma^*$. □

The following is a construction of an outer measure by an arbitrary set function γ defined on an arbitrary subcollection of sets $\mathcal{G} \subseteq \mathcal{P}(\Omega)$. As usual, we only assume that $\mathcal{G}$ contains the empty set and that γ, as a set function, is such that $\gamma(\emptyset) = 0$. This construction lies in the basis of the *Carathéodory extension* of the set function γ to a measure on σ-algebra $\Sigma(\mathcal{G})$.

For any subset $Q \subseteq \Omega$, denote by $\mathscr{C}_Q(\mathcal{G})$ the collection of all at most countable covers of set Q by elements of $\mathcal{G}$. [Unless there is another subcollection, besides $\mathcal{G}$, under consideration, we will for brevity drop $\mathcal{G}$ in $\mathscr{C}_Q(\mathcal{G})$.] Therefore, if $\mathscr{C}_Q \neq \emptyset$, for any $\{G_n\} \in \mathscr{C}_Q$, we have $Q \subseteq \bigcup_{n=1}^{\infty} G_n$.

2.8 Proposition. *The set function μ^* defined on $\mathcal{P}(\Omega)$ as*

$$\mu^*(Q) = \begin{cases} \inf\{\sum_{n=1}^{\infty}\gamma(G_n)\colon \{G_n\} \in \mathscr{C}_Q\}, & \mathscr{C}_Q \neq \emptyset \\ \infty, & \mathscr{C}_Q = \emptyset \end{cases} \tag{2.8}$$

is an outer measure.

Proof. We need to verify the above properties (2.1a - 2.1c) of μ^* as an outer measure:

a) Because $\emptyset \in \mathcal{G}$ and $\gamma(\emptyset) = 0,$ it follows that $\mu^*(\emptyset) = 0.$

b) We assume that both $\mu^*(A)$ and $\mu^*(B)$ are finite, because otherwise, the proof is obvious. If $A \subseteq B,\ \mathscr{C}_B \subseteq \mathscr{C}_A$ and then we can reach on $\mathscr{C}_A$ a possibly smaller limit inferior than that on $\mathscr{C}_B$. Therefore,

$$\mu^*(A) \leq \inf\{\textstyle\sum_{n=1}^{\infty}\gamma(G_n)\colon \{G_n\} \in \mathscr{C}_B\} = \mu^*(B).$$

c) Let $\{Q_n\} \subseteq \mathcal{P}(\Omega)$ and $Q = \bigcup\limits_{n=1}^{\infty} Q_n$. If for at least one n, $\mathscr{C}_{Q_n} = \emptyset,$ then also $\mathscr{C}_Q = \emptyset$ and subadditivity immediately follows. We assume that for all $n,\ \mathscr{C}_{Q_n} \neq \emptyset$ and choose an $\varepsilon > 0$. From

$$\mu^*(Q_i) = \inf\{\textstyle\sum_{n=1}^{\infty}\gamma(G_{in})\colon \{G_{in}\} \in \mathscr{C}_{Q_i}\},$$

and by the definition of a limit inferior, it follows that for $\varepsilon 2^{-i}$, there is a cover $\{G_{in},\ n = 1, 2, \dots\} \in \mathscr{C}_{Q_i}$ such that

$$\mu^*(Q_i) \leq \textstyle\sum_{n=1}^{\infty}\gamma(G_{in}) \leq \mu^*(Q_i) + \varepsilon 2^{-i}.$$

Now, clearly $\{G_{in},\ i, n = 1, 2, \dots\} \in \mathscr{C}_Q$. Thus,

$$\mu^*(Q) \leq \textstyle\sum_{i=1}^{\infty}\sum_{n=1}^{\infty}\gamma(G_{in}) \leq \sum_{i=1}^{\infty}\mu^*(Q_i) + \varepsilon,$$

which proves monotonicity. □

We call the couple $(\mathcal{G}, \gamma)$ [that is, a subset $\mathcal{G}$ of $\mathcal{P}(\Omega)$ and a set function γ on $\mathcal{G}$] a *formatter of outer measure* μ^* defined by (2.8). As has been shown, the formatter and, subsequently, the outer

measure, induced the σ-algebra Σ^*, on which μ^* was a complete measure.

When constructing a measure space $(\Omega, \Sigma^*, \mu_0^*)$ by $(\mathcal{G}, \gamma)$, the major goal is to extend γ from $\mathcal{G}$ to a measure, say μ, acting on the smallest σ-algebra $\Sigma(\mathcal{G})$ generated by $\mathcal{G}$. This can be achieved by restricting (Σ^*, μ_0^*) to $(\Sigma(\mathcal{G}), \mu)$, given that (Σ^*, μ_0^*) itself is an extension of $(\mathcal{G}, \gamma)$. The latter, however, is not guaranteed from the above construction unless we impose some restrictions upon the formatter $(\mathcal{G}, \gamma)$, for even though $(\mathcal{G}, \gamma)$ produces $(\Omega, \Sigma^*, \mu_0^*)$, $(\mathcal{G}, \gamma)$ need not have all elements μ^*-measurable. In other words, $\mathcal{G}$ need not be a subset of Σ^*. In addition, μ_0^* need not coincide with γ on $\mathcal{G}$. For example, if γ is an elementary content and $\mathcal{G}$ is a semi-ring, then, according to Problem 2.2, for each $G \in \mathcal{G}$, there is a cover $\{C_n\}$ of G such that $\Sigma_{n=1}^{\infty} C_n$ is a decomposition of G and

$$\mu^*(G) = \sum_{n=1}^{\infty} \gamma(C_n).$$

Hence, in order that $\mu^*(G) = \gamma(G)$, γ must be σ-additive on $\mathcal{G}$, which, in general, is not.

Consequently, we call (Σ^*, μ_0^*) [generated by $(\mathcal{G}, \gamma)$ in (2.8)] a *complete Carathéodory extension of* $(\mathcal{G}, \gamma)$ if $\mathcal{G} \subseteq \Sigma^*$ and $\mathrm{Res}_{\mathcal{G}} \mu_0^* = \gamma$. If (Σ^*, μ_0^*) is a Carathéodory extension of $(\mathcal{G}, \gamma)$, then the formatter $(\mathcal{G}, \gamma)$ is said to be *extendible* and the corresponding restriction of (Σ^*, μ_0^*) to $(\Sigma(\mathcal{G}), \mu)$ is referred to as a *Carathéodory extension of* $(\mathcal{G}, \gamma)$.

As mentioned above, one of the most important questions arises, as to what the formatter $(\mathcal{G}, \gamma)$ should really be to be extendible and, consequently, generate a Carathéodory extension. By now, we have a fairly large choice of systems of sets and set functions on them, ranging from semi-rings to σ-algebras and elementary contents to measures. The idea is, however, to select a possibly more rudimentary formatter $(\mathcal{G}, \gamma)$, which is tame and suited in most common practical applications and constructions and such that (Σ^*, μ_0^*) is an extension of $(\mathcal{G}, \gamma)$. In particular, this means that the elements of $\mathcal{G}$ have to be μ^*-measurable. The following theorem, which is a crucial

step in the whole extension procedure, infers that $(\mathcal{G}, \gamma)$ can be a ring and premeasure to serve as a reasonable and extendible formatter.

2.9 Theorem.

(i) *Let* $(\mathcal{G}, \gamma)$ *be a semi-ring and elementary content, respectively, in* Ω, *which produce the outer measure* μ^* *and* σ*-algebra* Σ^* *of* μ^**-measurable subsets of* Ω. *Then* $\mathcal{G} \subseteq \Sigma^*$.

(ii) *If, in addition,* γ *is* σ*-additive on* $\mathcal{G}$, *then* $\gamma = \mathrm{Res}_{\mathcal{G}}\mu^*$ *and therefore* (Σ^*, μ_0^*) *is an extension of* $(\mathcal{G}, \gamma)$.

In a nutshell, the formatter permits $\mathcal{G}$ to be as basic as a semi-ring, but the set function γ must be σ-additive. Now, if we extend $\mathcal{G}$ to the system that includes all finite unions of "rectangles" from $\mathcal{G}$, by Theorem 2.6, Chapter 4, we arrive at the ring $\mathcal{R}(\mathcal{G})$ to which γ is readily and naturally extendible, considering the utility of Lemma 2.5, Chapter 4 [according to which all figures of $\mathcal{R}(\mathcal{G})$ are unions of pairwise disjoint rectangles from $\mathcal{G}$] and Proposition 2.11 below, surely preserving its σ-additivity. Hence γ becomes a premeasure on $\mathcal{R}(\mathcal{G})$.

Proof.

(i) We have to show that $\mathcal{G} \subseteq \Sigma^*$, i.e., any element, $G \in \mathcal{G}$, μ^*-separates all subsets of Ω. Take any subset $Q \subseteq \Omega$ with $\mathscr{C}_Q \neq \emptyset$, because, otherwise, the proof would be trivial, and let $\mathcal{C} = \{C_n\}$ be any (countable) cover of Q from $\mathscr{C}_Q$. For a $G \in \mathcal{G}$, and $C_n \in \mathcal{C}$,

$$C_n = (C_n \cap G) + (C_n \cap G^c) = (C_n \cap G) + (C_n \backslash G). \quad (2.9)$$

Because $\mathcal{G}$ is a semi-ring, $C_n \cap G$ is an element of $\mathcal{G}$ and $C_n \backslash G$ can be represented as a finite union of pairwise disjoint elements of $\mathcal{G}$, say $\sum_{j=1}^{N_n} S_{j_n}$. Consequently, (2.9) can be rewritten as

$$C_n = (C_n \cap G) + \textstyle\sum_{j=1}^{N_n} S_{j_n}$$

and, by finite additivity of γ,

$$\gamma(C_n) = \gamma(C_n \cap G) + \textstyle\sum_{j=1}^{N_n} \gamma(S_{j_n}). \quad (2.9a)$$

Now, suppose $\sum_{n=1}^{\infty}\gamma(C_n) < \infty$. Then, summing up all equations in (2.9a) in n gives

$$\sum_{n=1}^{\infty}\gamma(C_n) = \sum_{n=1}^{\infty}\gamma(C_n \cap G) + \sum_{n=1}^{\infty}\gamma(S_n), \qquad (2.9b)$$

where $\{S_n\}$ is the reordered sequence $\{S_{j_n},\ \ j = 1, \dots, N_n,$ $n = 1, 2, \dots\}$.

As $Q = (Q \cap G) + (Q \cap G^c)$, obviously, $\{C_n \cap G\} \subseteq \mathcal{G}$ and $\{S_n\} \subseteq \mathcal{G}$ are covers of $Q \cap G$ and $Q \cap G^c$, respectively. Consequently,

$$\sum_{n=1}^{\infty}\gamma(C_n \cap G) \geq \mu^*(Q \cap G)$$

and

$$\sum_{n=1}^{\infty}\gamma(S_n) \geq \mu^*(Q \cap G^c),$$

and then by (2.9b),

$$\sum_{n=1}^{\infty}\gamma(C_n) \geq \mu^*(Q \cap G)\ + \mu^*(Q \cap G^c).$$

Because this inequality holds for every cover $\mathcal{C}$ of Q, it should also hold for the limit inferior to yield

$$\mu^*(Q) \geq \mu^*(Q \cap G)\ + \mu^*(Q \cap G^c). \qquad (2.9c)$$

If $\sum_{n=1}^{\infty}\gamma(C_n) = \infty$, then "equals symbol" in (2.9b) has to be replaced by " $\geq$ " to yield (2.9c) again. The inverse inequality is due to (2.3). Therefore, G separates all subsets of Ω and, consequently, we have $\mathcal{G} \subseteq \Sigma^*$.

(*ii*) By Problem 2.2, for each $G \in \mathcal{G}$, there is a cover $\{C_n\}$ of G such that $G = \sum_{n=1}^{\infty}C_n$ and $\mu^*(G) = \sum_{n=1}^{\infty}\gamma(C_n)$. Hence, if γ is σ-additive, μ^* coincides with γ on $\mathcal{G}$.

These two facts prove that (Σ^*, μ_0^*) is an extension of $(\mathcal{G}, \gamma)$. $\square$

2.10 Remarks.

(i) One should bear in mind that, although $(\mathcal{G}, \gamma)$ can serve as an extendible formatter for the outer measure μ^*, $\mathcal{G}$ is not really a generator for Σ^*, as the latter need not be the smallest σ-algebra containing $\mathcal{G}$. We want to make a clear distinction between these two terms. Recall that a family $\mathcal{G} \subseteq \mathcal{P}(\Omega)$ is said to be a *generator* of another family ($\mathcal{G} \subseteq$) $\mathcal{F}_0 \subseteq \mathcal{P}(\Omega)$ with a property P, if $\mathcal{F}_0$ is the intersection of all supercollections of $\mathcal{G}$ on each one of which property P holds. In our case, Σ^* will eventually contain the smallest σ-algebra $\Sigma = \Sigma(\mathcal{G})$ and, in general, μ^* needs to be further restricted to this σ-algebra. From Theorem 2.9, we conclude that any elementary content γ on a semi-ring $\mathcal{G}$, which is σ-additive, can be extended to a measure $\mu = \mathrm{Res}_{\Sigma(\mathcal{G})}\mu^*$ (acting on the smallest σ-algebra Σ generated by $\mathcal{G}$). In other words, if γ is a σ-additive elementary content on a semi-ring $\mathcal{G}$, then there exists at least one extension, namely, Carathéodory's extension.

(ii) From the proof of Theorem 2.9, it is obvious that a semi-ring of rectangles, with a σ-additive elementary content on this semi-ring, is one of the most economical systems suitable for the Carathéodory extension. However, it is often more preferable and common to start with premeasures on rings. In practice, one can first extend a semi-ring, with an elementary content, to the smallest ring with the content using the procedure proposed in Lemma 2.5 and Theorem 2.6, Chapter 4, and next Proposition 2.11.

(iii) Another reasonable question arises: in how many different ways can a formatter $(\mathcal{G}, \gamma)$ be extended to a measure on $\Sigma(\mathcal{G})$? Forthcoming Theorem 2.13 asserts that under some relatively minor restriction (Remark 2.12) made on set function γ, the uniqueness of Carathéodory's extension is guaranteed. □

We begin with one useful extension of an elementary content on a semi-ring to a content on the smallest ring containing the semi-ring.

2.11 Proposition. *There is exactly one content on $\mathcal{R}(\mathcal{S})$, which coincides with the elementary content on $\mathcal{S}$.* □

(See Problem 2.3.)

2.12 Remark. In Definition 1.1(vii) we introduced the notion of σ-finiteness of a set function. Sometimes it is more convenient to use another definition of σ-finiteness, which is equivalent to Definition 1.1(vii) for a large class of set functions. Namely, the condition to have a monotone increasing sequence $\{G_n\}\uparrow \Omega$ from $\mathcal{G}$, with $\gamma(G_n) < \infty$ for all n, can be replaced with the equivalent condition that there is a countable partition $\{\Omega_1, \Omega_2, \dots\} \subseteq \mathcal{G}$ of Ω ($= \Sigma_{n=1}^{\infty} \Omega_n$) such that $\gamma(\Omega_n) < \infty$ for all n. For instance, rings with contents clearly provide a basis for such equivalence. For a semi-ring with an elementary content, the first definition yields the second one, because we can arrange from $\{G_n\}\uparrow \Omega$ a countable decomposition; however, the converse, obviously, does not hold.

Another related feature we use in the sequel is the σ-finiteness of a particular set. Let (Ω, Σ, μ) be a measure space. A measurable set A is said to be *σ-finite* if $\mathrm{Res}_{\Sigma\cap A}\mu$ is σ-finite. □

2.13 Theorem. *Let $\mathcal{G}$ be a $\cap$-stable generator of the σ-algebra $\Sigma(\mathcal{G})$ in Ω such that $\mathcal{G}$ contains a monotone increasing sequence $\{B_n\}\uparrow \Omega$. Let μ_1 and μ_2 be two measures on $\Sigma(\mathcal{G})$, which are σ-finite on $\{B_n\}$ and coincide on $\mathcal{G}$. Then $\mu_1 = \mu_2$ on $\Sigma(\mathcal{G})$.*

Proof. Let $A \in \mathcal{G}$ such that $\mu_1(A) = \mu_2(A) < \infty$ and let $\mathcal{D}_A = \{B \in \Sigma\colon\ \mu_1(A \cap B) = \mu_2(A \cap B)\}$. We show that $\mathcal{D}_A$ is a Dynkin system:

$a)$ $A \in \mathcal{D}_A$ implies that $\Omega \in \mathcal{D}_A$.

$b)$ Let $D \in \mathcal{D}_A$. Then $A \cap D^c = A\backslash D = A\backslash(A \cap D)$, which implies that

$$\begin{aligned}\mu_1(A \cap D^c) &= \mu_1(A) - \mu_1(A \cap D)\\ &= \mu_2(A) - \mu_2(A \cap D) = \mu_2(A \cap D^c),\end{aligned}$$

and this leads to $D^c \in \mathcal{D}_A$.

$c)$ Let $\{D_n\}$ be a sequence of disjoint sets from $\mathcal{D}_A$. Then

$$\mu_1(A \cap \sum_{n=1}^{\infty} D_n) = \mu_1(\sum_{n=1}^{\infty} A \cap D_n) = \sum_{n=1}^{\infty} \mu_1(A \cap D_n)$$

$$= \sum_{n=1}^{\infty} \mu_2(A \cap D_n) = \mu_2(\sum_{n=1}^{\infty} A \cap D) = \mu_2(A \cap \sum_{n=1}^{\infty} D_n).$$

Hence $\sum_{n=1}^{\infty} D_n \in \mathcal{D}_A$, and therefore $\mathcal{D}_A$ is a Dynkin system. Obviously $\mathcal{G} \subseteq \mathcal{D}_A$. Consequently, $\mathcal{G} \subseteq \mathcal{D}(\mathcal{G}) \subseteq \mathcal{D}_A$. Also because $\mathcal{G}$ is $\cap$-stable, it follows that $\mathcal{D}(\mathcal{G})$ is a σ-algebra. Hence, we have

$$\mathcal{G} \subseteq \mathcal{D}(\mathcal{G}) = \Sigma(\mathcal{G}) \subseteq \mathcal{D}_A \subseteq \Sigma(\mathcal{G})$$

leading to

$$\mathcal{D}(\mathcal{G}) = \Sigma(\mathcal{G}) = \mathcal{D}_A.$$

In particular, we proved that $\forall B \in \Sigma(\mathcal{G})\ \mu_1(A \cap B) = \mu_2(A \cap B)$.

Now let $\{B_n\}$ be a monotone increasing sequence of sets from $\mathcal{G}$ convergent to Ω. Thus $\Sigma(\mathcal{G}) = \mathcal{D}_{B_n}$. Then $\forall n = 1, 2, \ldots,$ and $\forall B \in \Sigma(\mathcal{G})$,

$$\mu_1(B_n \cap B) = \mu_2(B_n \cap B).$$

Because $\{B_n \cap B\} \uparrow B$ and because $\mu_i(B \cap B_n) < \infty$, by Lemma 1.6,

$$\lim_{n \to \infty} \mu_1(B \cap B_n) = \lim_{n \to \infty} \mu_2(B \cap B_n)$$

$$= \mu_1(B) = \mu_2(B), \forall B \in \Sigma(\mathcal{G}). \qquad \square$$

Now, by means of Theorem 2.13 we easily deduce the following significant statement.

2.14 Corollary. *Let γ be a σ-finite and σ-additive elementary content on a semi-ring $\mathcal{G}$. Then a Carathéodory extension of γ to a measure on σ-algebra $\Sigma(\mathcal{G})$ is unique.* $\square$

The lemmas below are used for various purposes and, in particular, lead to a relationship between the completion $(\Omega, \overline{\Sigma}, \overline{\mu})$ of a measure space (Ω, Σ, μ) and the σ-algebra Σ^* of all μ^*-measurable sets.

2.15 Lemma. *Let* $(\Omega, \mathcal{G}, \gamma)$ *be an extendible formatter of the outer measure* μ^* *and* $\mathcal{G}_\sigma$ *the collection of all at most countable unions of elements from* $\mathcal{G}$. *Then, for each* $Q \subseteq \Omega$, *there is a set* $G_\sigma \in \mathcal{G}_\sigma$, *such that* $G_\sigma \supseteq Q$ *and*

$$\mu^*(G_\sigma) \leq \mu^*(Q) + \varepsilon. \tag{2.15}$$

Proof. Because μ^* is generated by $(\mathcal{G}, \gamma)$, we have

$$\mu^*(Q) = \begin{cases} \inf\{\sum_{n=1}^{\infty} \gamma(G_n) \colon \{G_n\} \in \mathscr{C}_Q\}, & \mathscr{C}_Q \neq \emptyset \\ \infty, & \mathscr{C}_Q = \emptyset. \end{cases} \tag{2.15a}$$

If $\mu^*(Q) = \infty$, then inequality (2.15) holds trivially. Suppose $\mu^*(Q) < \infty$. Then, by definition of a limit inferior and from (2.15a), for every $\varepsilon > 0$, there is $\{G_n\} \in \mathscr{C}_Q$ such that

$$\begin{aligned} &\mu^*(Q) + \varepsilon \\ &\qquad \geq \textstyle\sum_{n=1}^{\infty} \gamma(G_n) \geq \sum_{n=1}^{k} \gamma(G_n) \\ &\qquad = \textstyle\sum_{n=1}^{k} \mu^*(G_n) \geq \mu^*\left(\bigcup_{n=1}^{k} G_n \right). \end{aligned} \tag{2.15b}$$

Now, we make use of the fact that $(\mathcal{G}, \gamma)$ is an extendible formatter. This implies that not only $\mathcal{G} \subseteq \Sigma^*$, but also $\mathcal{G}_\sigma \subseteq \Sigma^*$. Since $\bigcup_{n=1}^{k} G_n$ is monotone increasing and $\mu^*\left(\bigcup_{n=1}^{k} G_n \right) < \infty$ for all k, by continuity from below (Lemma 1.6),

$$\lim_{k \to \infty} \mu^*\Big(\bigcup_{n=1}^{k} G_n \Big) = \mu^*\Big(\bigcup_{n=1}^{\infty} G_n \Big).$$

Passing to the limit in (2.15b), which holds true for all k, we prove (2.15) with $G_\sigma = \bigcup_{n=1}^{\infty} G_n$ being a pledged set. □

Lemma 2.16. *Let μ^* be an outer measure, Σ^* the σ-algebra of all μ^*-measurable sets, and A any subset of Ω. If there is a μ^*-measurable set B such that $B \supseteq A$ and $\mu^*(B\backslash A) = 0$, then $A \in \Sigma^*$.*

Proof. Because $B \in \Sigma^*$, it should μ^*-separate Q:

$$\mu^*(Q) = \mu^*(Q \cap B) + \mu^*(Q \cap B^c). \tag{2.16}$$

Now, because $A \subseteq B$, we can easily show that

$$Q \cap A^c = Q \cap B^c + Q \cap (B\backslash A). \tag{2.16a}$$

From $Q \cap (B\backslash A) \subseteq B\backslash A$, it follows that $\mu^*(Q \cap (B\backslash A)) = 0$. From (2.16a),

$$\mu^*(Q \cap A^c) = \mu^*(Q \cap B^c).$$

Consequently, we can replace $\mu^*(Q \cap B^c)$ in (2.16) by $\mu^*(Q \cap A^c)$. Finally, noticing that $Q \cap B \subseteq Q \cap A$, we have that

$$\mu^*(Q) = \mu^*(Q \cap B) + \mu^*(Q \cap B^c) \geq \mu^*(Q \cap A) + \mu^*(Q \cap A^c),$$

and this is the required inequality. □

Lemma 2.17. *Let μ^* be the outer measure generated by an extendible formatter $(\mathcal{G}, \gamma)$, Σ^* be the σ-algebra of all μ^*-measurable sets, μ_0^* be $\mathrm{Res}_{\Sigma^*}\mu^*$, and let $\Sigma(\mathcal{G})$ be the σ-algebra generated by $\mathcal{G}$. Then, for every $A^* \in \Sigma^*$ such that $\mu_0^*(A^*) < \infty$, there is a set $B \in \Sigma(\mathcal{G})$ with $B \supseteq A^*$ and $\mu_0^*(B\backslash A^*) = 0$.*

Proof. Because $\mu_0^*(A^*) < \infty$, $\mathscr{C}_Q \neq \emptyset$. From Lemma 2.15, for every $\varepsilon > 0$, say $\frac{1}{k}$, there is a $G_\sigma^k = \bigcup_{n=1}^{\infty} G_n^k \supseteq A^*$ such that $\mu_0^*(G_\sigma^k) \leq \mu_0^*(A^*) + \varepsilon$. The latter yields that $\mu_0^*(G_\sigma^k\backslash A^*) \leq \frac{1}{k}$. Obviously, $\bigcap_{k=1}^{m} G_\sigma^k$ is still a superset of A^* and in as much as $\bigcap_{k=1}^{m} G_\sigma^k \subseteq G_\sigma^m$, it follows that

$$\mu_0^*(D_m) \leq \mu_0^*(G_\sigma^m) \leq \tfrac{1}{m}, \tag{2.17}$$

where $D_m = \left(\bigcap_{k=1}^{m} G_\sigma^k \right) \backslash A^* \in \Sigma^*$. The sequence $\{D_m\}$ is clearly monotone nonincreasing and $\mu_0^*(D_1) < \infty$. Therefore, by continuity from above [see Theorem 1.7 (i)] of μ_0^* and because of (2.17),

$$n \underset{\rightarrow}{\lim}_{\infty} \mu_0^*(D_m) = \mu_0^* \left\{ \left(\bigcap_{k=1}^{\infty} G_\sigma^k \right) \backslash A^* \right\} = 0.$$

The set $\bigcap_{k=1}^{\infty} G_\sigma^k$ can serve as a set B asserted in the lemma, as it obviously meets the requirements regarding set B in the statement and we are done with the proof. □

Corollary 2.18. *Let μ^* be the outer measure generated by a σ-finite extendible formatter $(\mathcal{G}, \gamma)$, Σ^* be the σ-algebra of all μ^*-measurable sets, and let $\Sigma(\mathcal{G})$ be the σ-algebra generated by $\mathcal{G}$. Then, for every $A^* \in \Sigma^*$, there is a set $B \in \Sigma(\mathcal{G})$ with $B \supseteq A^*$ and $\mu^*(B \backslash A^*) = 0$.*

Proof. $(\mathcal{G}, \gamma)$ is σ-finite. Thus there is a partition $\{H_1, H_2, \dots\} \subseteq \mathcal{G}$ of Ω such that $\gamma(H_k) < \infty$. If $A^* \in \Sigma^*$, then

$$\{A_k^* = A^* \cap H_k,\ k = 1, 2, \dots\}$$

is a μ^*-measurable partition of A^*, with $\mu^*(A_k^*) < \infty$ for every k, and to each of which we can apply Lemma 2.17 and have a set $B_k \in \Sigma(\mathcal{G})$, with $B_k \supseteq A_k^*$ and $\mu^*(B_k \backslash A_k^*) = 0$.

Notice that because

$$\begin{aligned} \left(\bigcup_{k=1}^{\infty} B_k \right) \backslash \left(\sum_{n=1}^{\infty} A_n^* \right) &= \left(\bigcup_{k=1}^{\infty} B_k \right) \cap \left(\bigcap_{n=1}^{\infty} (A_n^*)^c \right) \\ &= \bigcup_{k=1}^{\infty} \left\{ B_k \cap \left(\bigcap_{n=1}^{\infty} (A_n^*)^c \right) \right\} \\ &\subseteq \bigcup_{k=1}^{\infty} \{ B_k \cap (A_k^*)^c \}, \end{aligned}$$

it holds true that

$$\mu^*\left[\left(\bigcup_{k=1}^{\infty} B_k\right)\backslash\left(\sum_{n=1}^{\infty} A_n^*\right)\right]$$

$$\leq \mu^*\left[\bigcup_{k=1}^{\infty}(B_k\backslash A_k^*)\right] \leq \sum_{n=1}^{\infty}\mu^*(B_n\backslash A_n^*) = 0.$$

The statement follows after letting $B = \bigcup_{k=1}^{\infty} B_k$ $(\in \Sigma(\mathcal{G}))$. □

Now, with the aid of the above propositions, we can finally answer the question about the relationship between the completion $(\Omega, \overline{\Sigma}, \overline{\mu})$ of a measure space (Ω, Σ, μ) and the measure space $(\Omega, \Sigma^*, \mu_0^*)$ (which is also complete, as we learned it).

2.19. Theorem. *Let $(\mathcal{G}, \gamma)$ be an extendible formatter for $(\Omega, \Sigma^*, \mu_0^*)$ and a generator for the measure space $(\Omega, \Sigma = \sigma(\mathcal{G}), \mu = \mathrm{Res}_{\Sigma}\mu^*)$ whose completion is $(\Omega, \overline{\Sigma}, \overline{\mu})$.*

(i) *Then, $\overline{\Sigma} \subseteq \Sigma^*$.*

(ii) *If $(\mathcal{G}, \gamma)$ is σ-finite, then $\overline{\Sigma} = \Sigma^*$ and $\overline{\mu} = \mu_0^*$.*

Proof.

(i) Obviously, $\overline{\Sigma} \subseteq \Sigma^*$ if and only if any element $\overline{A}$ of $\overline{\Sigma}$ is of the form $A \cup N$, where $A \subseteq \Sigma$, N is μ-negligible, and $\overline{A}$ is μ^*-measurable. According to Lemma 2.16, $A \cup N$ would be μ^*-measurable, if there is a μ^*-measurable set B such that $B \supseteq A \cup N$ and $\mu^*(B\backslash(A \cup N)) = 0$. By Definition 2.5$(i)$ of a μ-negligible set, N must have a Σ-measurable μ-null superset, say N_0. [Note that even though, by Problem 2.10, $\mu^*(N) = 0$ and $\mu^*(A \cup N) = \mu^*(A)$, this does not warrant $A \cup N \in \Sigma^*$.] Because $A \cup N_0$ is a superset of $A \cup N$ and, by Problem 2.11, $(A \cup N_0)\backslash(A \cup N)$ is a μ^*-null set, $B = A \cup N_0$ meets all prerequisites of Lemma 2.16, which makes $A \cup N$ indeed μ^*-measurable. This proves part (i) of the theorem.

(ii) Because of part (i), we need to show that $\Sigma^* \subseteq \overline{\Sigma}$, that is, each A^* can be represented as the union of a μ-measurable set and μ-negligible set. By Problem 2.12, for any $A^* \in \Sigma^*$, there is a Σ-measurable subset B of A^* such that $\mu^*(A^*\backslash B) = 0$. Obviously, A^* can be decomposed as B and μ^*-measurable null set $A^* \cap B^c$. It only remains to show that $A^* \cap B^c$ is μ-negligible.

By Corollary 2.18, for A^*, there is a set $C \in \Sigma$ such that $C \supseteq A^*$ and $\mu^*(C \backslash A^*) = 0$. The set-difference $C \backslash B = (C \backslash A^*) + (A^* \backslash B)$, as the union of two μ^*-null sets, is a μ^*-null set, therefore, a μ-null set (as $C \backslash B \in \Sigma$). This proves that $A^* \cap B^c$ is μ-negligible.

Now, we show that $\overline{\mu} = \mu_0^*$. (Recall that they are equal on Σ.) Because $\Sigma = \Sigma^*$, $A^* = A \cup N$, where $A \in \Sigma$ and N is μ-negligible, and

$$\overline{\mu}(A^*) = \mu(A) = \mu^*(A). \tag{2.19}$$

On the other hand, there is a μ-null superset of N to yield $\mu^*(N) = 0$ due to monotonicity of μ^*. Finally, from the inequalities

$$\mu^*(A^*) \leq \mu^*(A) + \mu^*(N) = \mu^*(A)$$

and

$$\mu^*(A^* = A \cup N) \geq \mu^*(A),$$

it follows that $\mu^*(A^*) = \mu^*(A)$ and this, along with (2.19), yields that $\overline{\mu}(A^*) = \mu^*(A^*)$ for each $A^* \in \Sigma^* = \overline{\Sigma}$. □

Example 2.20. If (Ω, Σ, μ) is a probability space, it follows from Theorem 2.19 that the completion of (Σ, μ) coincides with (Σ^*, μ_0^*) induced by (Σ, μ) or by a "smaller generator" $(\mathcal{G}, \gamma)$ of (Σ, μ). □

A noteworthy question arises: if we have a semi-ring and σ-additive elementary content, would it make any difference if we first extend them to the smallest ring and premeasure, according to Proposition 2.11, and then use the Carathéodory extension to arrive at the smallest σ-algebra and a measure on it, or alternatively, apply the Carathéodory extension directly to that semi-ring and σ-additive elementary content. The same question applies, say, to a ring with a premeasure and the generated σ-algebra with a measure. The difference, if any, can apparently be manifested by the emergence of two outer measures, induced by a formatter and its extension.

2.21 Theorem. *Let $(\Omega, \mathcal{G}, \gamma_0)$ be an extendible formatter of outer measure μ^* and σ-algebra Σ^* of μ^*-measurable sets and let $(\mathcal{E} = \mathcal{E}(\mathcal{G}), \gamma)$ be an extension of $(\mathcal{G}, \gamma_0)$ and an extendible formatter of outer measure ν^* and $\mathcal{B}^*$, such that $\mathcal{E} \subseteq \Sigma^*$ and $\gamma = \mathrm{Res}_{\mathcal{E}} \mu^*$. Then, $\nu^* = \mu^*$ on $\mathcal{P}(\Omega)$ and $\Sigma^* = \mathcal{B}^*$.*

Proof. Let $Q \subseteq \Omega$. Because $\mathscr{C}_Q(\mathcal{G}) \subseteq \mathscr{C}_Q(\mathcal{E})$, obviously

$$\nu^*(Q) \leq \mu^*(Q), \tag{2.21}$$

which yields the equation $\nu^* = \mu^*$ on a subcollection of sets $Q \in \mathcal{P}(\Omega)$ with $\nu^*(Q) = \infty$. Suppose $\nu^*(Q) < \infty$. Then, for every $\frac{\varepsilon}{2} > 0$, there is a cover $\{E_n\} \in \mathscr{C}_Q(\mathcal{E})$ with

$$\textstyle\sum_{n=1}^{\infty} \gamma(E_n) \leq \nu^*(Q) + \frac{\varepsilon}{2}. \tag{2.21a}$$

Because $\gamma = \mu^*$ on $\mathcal{E}$ and $\gamma(E_n) = \mu^*(E_n) < \infty$, for each $\varepsilon 2^{-n-1} > 0$, there is a cover $\{G_{nk}, k = 1,2,\dots\} \in \mathscr{C}_{E_n}(\mathcal{G})$, such that

$$\begin{aligned} \textstyle\sum_{k=1}^{\infty} \gamma_0(G_{nk}) &\leq \mu^*(E_n) + \varepsilon 2^{-n-1} \\ &= \gamma(E_n) + \varepsilon 2^{-n-1}. \end{aligned} \tag{2.21b}$$

Because $\{G_{nk}, n,k = 1,2,\dots\} \in \mathscr{C}_Q(\mathcal{G})$, from (2.21a) and (2.21b),

$$\mu^*(Q) \leq \textstyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \gamma_0(G_{nk}) \leq \nu^*(Q) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2}. \tag{2.21c}$$

Finally, taking in (2.21c) $\varepsilon = \frac{1}{m}$ leads to the inverse of inequality (2.21) and proves that $\mu^* = \nu^*$ on $\mathcal{P}(\Omega)$. The outer measure is the mere generator of the σ-algebra of separating sets. Consequently, $\mu^* = \nu^*$ yields that $\Sigma^* = \mathcal{B}^*$, which completes the proof. □

An important consequence of Theorem 2.21 is the following.

2.22 Corollary. *Let $(\mathcal{S}, \gamma_0)$ be a semi-ring, σ-additive and σ-finite elementary content in Ω, and let $(\mathcal{E} = \mathcal{E}(\mathcal{S}), \gamma)$ be an extension of $(\mathcal{S}, \gamma_0)$ such that $\mathcal{E} \subseteq \Sigma(\mathcal{S})$. Suppose $(\Sigma_s = \Sigma(\mathcal{S}), \mu_s)$ and $(\Sigma_e = \Sigma(\mathcal{E}), \mu_e)$ are the Carathéodory extensions with their respective outer measures μ_s^* and μ_e^* and σ-algebras Σ_s^* and Σ_e^* of measurable sets. Then, the following hold true.*

1) $\Sigma = \Sigma_s = \Sigma_e$.

2) $\mu = \mu_s = \mu_e$ on Σ.

3) $\mu^* = \mu_s^* = \mu_e^*$ on $\mathcal{P}(\Omega)$.

4) $\Sigma^* = \Sigma_s^* = \Sigma_e^*$.

Proof.

1) From $\mathcal{E} \subseteq \Sigma_s$ we have $\Sigma_e \subseteq \Sigma_s$. From $\mathcal{S} \subseteq \mathcal{E} \subseteq \Sigma_e$ it follows that $\Sigma_s \subseteq \Sigma_e$.

2) Now measures μ_s and μ_e act on the same σ-algebra Σ and coincide on semi-ring $\mathcal{S}$. In as much as γ_0 is σ-finite on $\mathcal{S}$, by Corollary 2.14, $\mu_s = \mu_e$ on Σ.

3) With

$$\mathcal{S} \subseteq \mathcal{E} \subseteq \Sigma_s \subseteq \Sigma_s^*, \ \mu = \mathrm{Res}_{\Sigma}\mu_e^* = \mathrm{Res}_{\Sigma}\mu_s^*$$

and, consequently,

$$\gamma = \mathrm{Res}_{\mathcal{E}}\mu_s^* = \mathrm{Res}_{\mathcal{E}}\mu_e^*,$$

we meet all conditions of Theorem 2.21 to have $\mu_e^* = \mu_s^* = \mu^*$.

4) $\Sigma^* = \Sigma_s^* = \Sigma_e^*$ also by Theorem 2.21. □

For instance, $\mathcal{E}$ can be a ring generated by $\mathcal{S}$ and γ, the extension of the elementary content γ_0 in accordance with Proposition 2.11; or $\mathcal{E}$ can be an algebra with γ as a premeasure or $\mathcal{E}$ can even be the σ-algebra $\Sigma(\mathcal{S})$. In particular, it follows that, once the Carathéodory extension from $(\mathcal{S}, \gamma_0)$ to $(\Sigma(\mathcal{S}), \mu)$ is rendered, another Carathéodory extension of $(\mathcal{E},\gamma)$ would be redundant.

Another consequence of Theorem 2.21 is the uniqueness of outer measures generated by measures.

Corollary 2.23. *Let μ be a measure on a σ-algebra Σ, which produces the outer measure μ^* with σ-algebra Σ^* of measurable sets. If there is another outer measure $\widetilde{\mu}^*$, then $\widetilde{\mu}^* = \mu^*$ on $\mathcal{P}(\Omega)$ and $\Sigma^* = \widetilde{\Sigma}^*$.*

Proof. This is a direct application of Theorem 2.21 with the following identification of the above characteristics:

1) Let $\widetilde{\mu}$ be a measure on Σ such that $\widetilde{\mu} = \mu$. Then $\widetilde{\mu}$ can serve as an extension of (Σ, μ).

2) $\Sigma \subseteq \Sigma^*$.

3) $\mu = \widetilde{\mu} = \text{Res}_{\Sigma}\mu^*$. □

Remark 2.24. Corollary 2.23 is useful in various applications of Carathéodory's extension. Suppose γ_1 and γ_2 are two elementary contents coinciding on a σ-finite semi-ring $\mathcal{S}$ (i.e. they are σ-finite on $\mathcal{S}$). By Corollary 2.14, their respective Carathéodory extensions μ_1 and μ_2 must coincide on $\Sigma(\mathcal{S})$. Let μ_1^* and μ_2^* be the corresponding outer measures, according to Corollary 2.22, produced by γ_1 and γ_2 or μ_1 and μ_2 (regardless). By Corollary 2.23, $\mu_1^* = \mu_2^*$ on $\mathcal{P}(\Omega)$ and $\Sigma_1^* = \Sigma_2^*$. □

As in Theorem 2.21, by comparing two measures generated by a set function acting on a collection of sets and their extension, we ended up comparing two corresponding produced outer measures. It seems to be reasonable to raise another question: what if an outer measure will produce another outer measure? Would this make any difference? More specifically, can the restriction μ_0^* of an outer measure μ^* on Σ^* become a formatter of another, different from μ^*, outer measure? Note that this is a different scenario from one considered in Theorem 2.21, because here μ^* is not supposed to be generated by a formatter and it "acts on its own." The following example shows us this distinction.

2.25 Example. Consider $\mathcal{P}(\Omega)$, μ^*, Σ^*, and μ_0^* in Example 2.4 (i):

$$\Omega = \{a, b, c\},\ A = \{a\},\ A^c = \{b, c\},\ P = \{b\},$$

$$Q = \{c\},\ R = \{a, b\},\ S = \{a, c\},$$

$$\mu^*(\emptyset) = 0,\ \mu^*(\Omega) = 4,\ \mu^*(A) = 1,$$

$$\mu^*(A^c) = \mu^*(R) = \mu^*(S) = 3,\ \mu^*(P) = \mu^*(Q) = 2,$$

$$\Sigma^* = \{\emptyset, \Omega, A, A^c\},\ \text{ and }\ \mu_0^* = \text{Res}_{\Sigma^*}\mu^*.$$

Then, generate the outer measure ν^* by (Σ^*, μ_0^*). So, we have:

$$\mu^* = \nu^* \text{ on } \Sigma^* \text{ and } \nu^*(P) = \mu^*(A^c) = 3\ (> \mu^*(P) = 2),$$

$$\nu^*(Q) = \mu^*(A^c) = 3\ (> \mu^*(Q) = 2),$$

$$\nu^*(R) = \mu^*(\Omega) = 4\ (> \mu^*(R) = 3),$$

$$\nu^*(S) = \mu^*(\Omega) = 4\ (> \mu^*(S) = 3).$$

As we see, in most cases ν^* is strictly greater than μ^* on $\mathcal{P}(\Omega)$. □

According to Problem 2.1, if ν^* is an outer measure induced by μ_0^*, then $\mu^* \leq \nu^*$ always holds on $\mathcal{P}(\Omega)$. That $\mu^* = \nu^*$ requires some restrictions, such as those in the proposition below.

2.26 Proposition. *Let μ^* be an outer measure on $\mathcal{P}(\Omega)$, $\mu_0^* = \text{Res}_{\Sigma^*}\mu^*$ and ν^* be the outer measure produced by (Σ^*, μ_0^*). The equation $\mu^* = \nu^*$ holds true on $\mathcal{P}(\Omega)$ if and only if for every $Q \in \mathcal{P}(\Omega)$, there exists a set $A^* \in \Sigma^*$ such that $A^* \supseteq Q$ and $\mu^*(A) = \mu^*(Q)$.* □

(See Problem 2.17.)

2.27 Remark. If μ^* is generated by an extendible formatter $(\mathcal{G}, \gamma)$, then clearly $\mu^* = \nu^*$, due to Theorem 2.21, as (μ_0^*, Σ^*) can serve as an extension of $(\mathcal{G}, \gamma)$. Alternatively, if $Q \in \mathcal{P}(\Omega)$, according to Lemma 2.15, for each positive ε, there is a set $G_\sigma \in \mathcal{G}_\sigma$ (a collection of all countable unions of elements from $\mathcal{G}$) such that $G_\sigma \supseteq Q$ and $\mu^*(G_\sigma) \leq \mu^*(Q) + \varepsilon$. We assume that ν^* is the outer measure generated by μ_0^*. Because $\mu^* = \nu^*$ on $\mathcal{P}(\Omega)$ and $G_\sigma \in \Sigma^*$, we have $\mu^*(G_\sigma) = \nu^*(G_\sigma)$ and, by monotonicity, $\nu^*(G_\sigma) \geq \nu^*(Q)$. Thus, we have

$$\nu^*(Q) \leq \mu^*(G_\sigma) = \nu^*(G_\sigma) \leq \mu^*(Q) + \varepsilon,$$

which yields $\nu^*(Q) \leq \mu^*(Q)$. The inverse inequality is due to Problem 2.1. □

2.28 Theorem. *Let (Ω, Σ, μ) be a measure space such that $\Sigma = \Sigma(\mathcal{S})$ with $\mathcal{S}$ being a semi-ring, and μ be σ-finite on $\mathcal{S}$. Then, given $A \in \Sigma(\mathcal{S})$ and $\varepsilon > 0$, there is a disjoint countable cover*

$\{S_n\} \subseteq \mathcal{S}$ of A that "approximates" A, that is, such that $A \subseteq \Sigma_{n=1}^{\infty} S_n$ and $\mu((\Sigma_{n=1}^{\infty} S_n) \backslash A) < \varepsilon$.

Proof. Let $\gamma = \text{Res}_{\mathcal{S}}\mu$ and μ^* be the outer measure produced by $(\mathcal{S}, \gamma)$. Then, μ is the unique Carathéodory extension of γ from $\mathcal{S}$ to $\Sigma(\mathcal{S})$, according to Corollary 2.14, and $\mu = \text{Res}_{\Sigma}\mu^*$.

Case 1. Let $\mu(A) = \mu^*(A) < \infty$.

Then, by (2.8) (of Proposition 2.8), for each $\varepsilon > 0$, there is a sequence $\{G_n\} \in \mathscr{G}_A$ such that

$$\sum_{n=1}^{\infty} \gamma(G_n) = \sum_{n=1}^{\infty} \mu(G_n) < \mu(A) + \varepsilon.$$

Since $\mu\left(\bigcup_{n=1}^{\infty} G_n\right) \leq \sum_{n=1}^{\infty} \mu(G_n)$, we have that

$$\mu\left(\left\{\bigcup_{n=1}^{\infty} G_n\right\} \backslash A\right) < \varepsilon.$$

Case 2. $\mu(A)$ is arbitrary.

By σ-finiteness of μ on $\mathcal{S}$, there is at most a countable decomposition $\sum_{k=1}^{\infty} \Omega_k$ of Ω by $\{\Omega_k\} \subseteq \mathcal{S}$ such that $\mu(\Omega_k) < \infty$ and hence $\mu(A \cap \Omega_k) < \infty$. Now, we apply the above arguments to $A \cap \Omega_k$ and $\frac{\varepsilon}{2^k}$. Hence, there is a sequence $\{G_{nk}\} \subseteq \mathcal{S} \cap \Omega_k$ such that

$$\mu\left(\left(\bigcup_{n=1}^{\infty} G_{nk}\right) \backslash (A \cap \Omega_k)\right) < \tfrac{\varepsilon}{2^k}.$$

This leads to

$$\mu\left(\sum_{k=1}^{\infty}\left(\bigcup_{n=1}^{\infty} G_{nk}\right) \backslash (A \cap \Omega_k)\right) < \varepsilon$$

and thus to

$$= \mu\left(\left(\bigcup_{n=1}^{\infty} G_n\right) \backslash A\right) < \varepsilon,$$

where $G_n := \sum_{k=1}^{\infty} G_{nk}$.

Finally, it remains to form a disjoint sequence of semi-ring sets as stated in the theorem. The latter can be rendered in the same way as in Lemma 2.5 of Chapter 4. □

2.29 Corollary. *If under the condition of Theorem* 2.28, $\mu(A) < \infty$, *then* A *can be approximated by just a finite tuple of disjoint semi-ring sets.*

Proof. Because $\mu(A) < \infty$, by Case 1 of Theorem 2.28, so is $\mu(\Sigma_{n=1}^{\infty} G_n) < \infty$. Then, by continuity from above, for each $\varepsilon > 0$, there is an N such that

$$\mu(\textstyle\sum_{k=n+1}^{\infty} G_k) < \varepsilon, \text{ for all } n \geq N$$

thereby leading to

$$\mu(A \Delta (\textstyle\sum_{k=1}^{n} S_k)) < 2\varepsilon. \qquad \square$$

PROBLEMS

2.1 Let μ^* be an outer measure on $\mathcal{P}(\Omega)$, $\mu_0^* = \text{Res}_{\Sigma^*}\mu^*$, and ν^* be the outer measure induced by μ_0^*. Show that $\mu^* \leq \nu^*$ on $\mathcal{P}(\Omega)$.

2.2 Let $(\mathcal{G}, \gamma)$ be a formatter of the outer measure μ^* defined by (2.8). Show that if γ is an elementary content and $\mathcal{G}$ is a semi-ring, then for each $G \in \mathcal{G}$, there is a cover $\{C_n\}$ of G such that $G = \sum_{n=1}^{\infty} C_n$ and $\mu^*(G) = \sum_{n=1}^{\infty} \gamma(C_n)$.

2.3 Prove Proposition 2.11. *There is exactly one content on* $\mathcal{R}(\mathcal{S})$, *which coincides with the elementary content on* $\mathcal{S}$

2.4 Let μ be a finite measure on (Ω, Σ) and let $\widetilde{\Sigma}$ be any sub-collection of Σ. Show that, for any fixed subset $Q \subseteq \Omega$, it is true that

$$\inf\{\mu(A)\colon Q \subseteq A\} = \mu(\Omega) - \sup\{\mu(A^c)\colon Q \subseteq A\}.$$

2.5 Show that the original definition of σ-finiteness $1.1(vii)$ implies the second definition of σ-finiteness for semi-rings and elementary contents mentioned in Remark 2.12.

2.6 Let μ^* be an outer measure on $\mathcal{P}(\Omega)$ and $\{A_n\}$ a sequence of disjoint μ^*-measurable sets. Show that for any $Q \subseteq \Omega$,

$$\mu^*(Q \cap \sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu^*(Q \cap A_n).$$

2.7 Let $N \in \mathcal{N}_\mu$ (i.e., a μ-null set) and let $B \in \Sigma$. Show that $\mu(N \cup B) = \mu(B \backslash N) = \mu(B)$.

2.8 Show that $\overline{\Sigma}$ defined in Definition $2.5(ii)$ is a σ-algebra, $\overline{\mu}$ is a measure, that this extension does not depend upon representations of sets of $\overline{\Sigma}$, and that $(\Omega, \overline{\Sigma}, \overline{\mu})$ is complete.

2.9 Show that the measure space defined in Example $1.2(iv)$ is complete.

2.10 Let μ^* be an outer measure on $\mathcal{P}(\Omega)$ and $N \subseteq \Omega$ be such that $\mu^*(N) = 0$. Show that for any subset $Q \subseteq \Omega$, $\mu^*(Q \cup N) = \mu^*(Q)$.

2.11 Show that $(A \cup N_0) \backslash (A \cup N)$ in part (i) of Theorem 2.19 is a μ^*-null set.

2.12 Let μ^* be the outer measure generated by an extendible σ-finite formatter $(\mathcal{G}, \gamma)$, Σ^* be the σ-algebra of all μ^*-measurable sets, and let $\Sigma(\mathcal{G})$ be the σ-algebra generated by $\mathcal{G}$. Show that for every $A^* \in \Sigma^*$, there is a set $B \in \Sigma(\mathcal{G})$ with $B \subseteq A^*$ and $\mu^*(A^* \backslash B) = 0$.

2.13 Let $(\Omega, \Sigma_0, \mu_0)$ be a completion of a measure space (Ω, Σ, μ). Define for each $A \subseteq \Omega$ $\overline{\mu}(A) = \sup\{\mu(B)\colon\ B \in \Sigma,\ B \subseteq A\}$ and $\underline{\mu}(A) = \inf\{\mu(B)\colon B \in \Sigma,\ A \subseteq B\}$. Show that if

$a)$ $A \in \Sigma_0$, then $\overline{\mu}(A) = \underline{\mu}(A) = \mu_0(A)$,

$b)$ $\overline{\mu}(A) = \underline{\mu}(A) < \infty$, then $A \in \Sigma_0$.

2.14 Let Σ be a σ-algebra in Ω and let $a \in \Omega$. Show that for $\{a\} \in \Sigma$, the measure space $(\Omega, \Sigma, \varepsilon_a)$ is complete if and only if $\Sigma = \mathcal{P}(\Omega)$.

2.15 (Generalization of Problem 1.12.) Let (Ω, Σ, μ) be a measure space, μ be σ-finite, $\mathcal{G}_\infty = \{G \in \Sigma : \mu(G) < \infty\}$, and $\Sigma_\infty = \{Q \subseteq \Omega : Q \cap G \in \Sigma, \forall G \in \mathcal{G}_\infty\}$. Show that $\Sigma_\infty = \Sigma$.

2.16 Let (Ω, Σ, μ) be a measure space, $\mathcal{G}_\infty = \{G \in \Sigma : \mu(G) < \infty\}$, and $\Sigma_\infty = \{Q \subseteq \Omega : Q \cap G \in \Sigma, \forall G \in \mathcal{G}_\infty\}$. Define the set function μ_∞ on Σ_∞ as

$$\mu_\infty(Q) = \begin{cases} \mu(Q), & Q \in \Sigma \\ \infty, & Q \notin \Sigma. \end{cases}$$

Problem 1.13 claimed that μ_∞ was a measure on Σ_∞. Show that if the measure space (Ω, Σ, μ) is complete, then so is μ_∞.

2.17 Prove Proposition 2.26. *Let μ^* be an outer measure on $\mathcal{P}(\Omega)$, $\mu_0^* = \mathrm{Res}_{\Sigma^*}\mu^*$, and ν^* be the outer measure produced by (Σ^*, μ_0^*). The equation $\mu^* = \nu^*$ holds true on $\mathcal{P}(\Omega)$ if and only if for every $Q \in \mathcal{P}(\Omega)$, there exists a set $A^* \in \Sigma^*$ such that $A^* \supseteq Q$ and $\mu^*(A) = \mu^*(Q)$.*

2.18 Let μ^* be the outer measure generated by an extendible formatter $(\mathcal{G}, \gamma)$ on a non-empty set Ω, Σ^* be the σ-algebra of all μ^*-measurable sets, and $\Sigma(\mathcal{G})$ be the σ-algebra generated by $\mathcal{G}$. Show that a subset $N \subseteq \Omega$ is negligible if and only if $\mu^*(N) = 0$.

NEW TERMS:

Carathéodory's extension procedure, 283
outer measure, 283, 284
monotonicity of outer measure, 283, 285
σ-subadditivity of outer measure, 283, 285
μ^*-measurable set, 283, 285
Σ^*-σ-algebra, 283
μ^*-separability, 283, 285
μ_0^*-measure, 283
complete measure, 283
formatter of the outer measure, 284, 292
extension of the formatter to a measurable space, 284
extendible formatter, 284, 293
complete Carathéodory's extension of a formatter, 284, 293
Carathéodory's extension of a formatter, 284, 293
extendibility conditions for the formatter, 284
Carathéodory, Constantin, 285
Carathéodory's extension theorem, 285
μ-null (null) set, 290
null (μ-null) set, 290
$\mathcal{N}_\mu$-set, 290
μ-negligible (negligible) set, 290
negligible (μ-negligible) set, 290
complete measure space, 290
completion of a measure, 290
completion of a measure space, 290
restriction of outer measure to Σ^*-algebra, 291
extendibility of a formatter, criterion of, 294
σ-finiteness of a set, 297
Carathéodory's extension, uniqueness of, 297, 298
outer measure, uniqueness of, 305, 307

3. LEBESGUE AND LEBESGUE - STIELTJES MEASURES

In this section, we utilize the results of the previous section for the construction of Lebesgue and Lebesgue - Stieltjes measures. We have learned that to warrant the Carathéodory extension, a given formatter should be at least a semi-ring and σ-additive elementary content, which applies to some important special cases of formatters in Euclidean spaces. In Theorem 3.1 below, we show that the Lebesgue content is σ-additive on the ring $\mathcal{R}(\mathbb{R}^n)$, which implies that the Lebesgue elementary content is also σ-additive on the semi-ring of half-open intervals. Although it is possible to prove this statement directly (with no prior extension and Ø-continuity arguments, as in Theorem 3.1), in order to exploit the equivalence of Ø-continuity and σ-additivity, we prefer to start with the extension of the Lebesgue content on the ring $\mathcal{R}(\mathbb{R}^n)$ to the Lebesgue premeasure.

Theorem 3.1. *The Lebesgue content λ_c on the ring $\mathcal{R}(\mathbb{R}^n)$ is σ-additive, that is, a premeasure.*

Proof. Because the Lebesgue content λ_c is finite on $\mathcal{R}$, by Proposition 1.7 (ii), λ_c would be a premeasure if it were Ø-continuous. We use the equivalent version of Ø-continuity:

> *For every monotone decreasing sequence $\{A_k\}\downarrow \subseteq \mathcal{R}$ with $\lambda_c(A_1) < \infty$, the assumption that $\lim_{k\to\infty} \lambda_c(A_k)$* (which clearly exists) *is strictly positive must yield that $\bigcap_{k=1}^{\infty} A_k \neq \emptyset$.*

Let $\{A_k\}$ be any such monotone decreasing sequence with

$$\varepsilon = \lim_{k\to\infty} \lambda_c(A_k) > 0. \tag{3.1}$$

It is readily seen that (3.1) implies that for each k, $A_k \neq \emptyset$, and therefore, by Cantor's Theorem 5.4, Chapter 2, $\bigcap_{k=1}^{\infty} \overline{A}_k \neq \emptyset$. However, the nonempty intersection of the closures of A_k's does not

imply that the intersection $\bigcap_{k=1}^{\infty} A_k \neq \emptyset$ either. To overcome this predicament we construct a subsequence of compact subsets of A_k's whose intersection is nonempty.

Now, in as much as A_k's $\in \mathcal{R}$, each A_k can be represented as a finite union of disjoint half-open rectangles, say $\Sigma_{s=1}^{r} P_s$ (for brevity let us drop index k) such that $\lambda_c(P_s) > 0$. Then for each value of ε and for every P_s, there is a half-open rectangle Π_s whose closure $\overline{\Pi}_s$ is a proper subset of P_s and such that

$$\lambda_c(P_s) \leq \lambda_c(\Pi_s) + \tfrac{2^{-k}}{2^s - 2^{s-r}}\varepsilon. \tag{3.1a}$$

Bound (3.1a) yields that

$$\lambda_c(A_k) \leq \lambda_c(B_k) + \tfrac{1}{2^k}\varepsilon, \tag{3.1b}$$

where $B_k = \sum_{s=1}^{r} \Pi_s$. Obviously, $\overline{B}_k \subseteq A_k$. It seems as if we are done with the sequence $\{B_k\}$. However, the claim that $\bigcap_{k=1}^{\infty} \overline{B}_k \neq \emptyset$ is unwarranted, since $\{B_k\}$ need not be monotone decreasing. Therefore, we introduce

$$C_k = \bigcap_{j=1}^{k} B_j,$$

which forms a monotone nonincreasing sequence of sets termwise dominated by $\{A_k\}$. Now, we need to show that $C_k \neq \emptyset$. In fact, we will be able to prove a much stronger statement that $\lambda_c(C_k) > 0$ for all k. Namely, we prove that

$$\lambda_c(C_k) \geq \lambda_c(A_k) - \varepsilon\big(1 - \tfrac{1}{2^k}\big), \tag{3.1c}$$

which, because of $\lambda_c(A_k) \geq \varepsilon$, would yield the needed inequality:

$$\lambda_c(C_k) \geq \varepsilon\tfrac{1}{2^k}. \tag{3.1d}$$

We prove (3.1c) by induction. For $k = 1$, (3.1c) holds true, because from (3.1b),

$$\lambda_c(C_1) = \lambda_c(B_1) \geq \lambda_c(A_1) - \tfrac{1}{2}\varepsilon.$$

Now we assume that (3.1c) holds for some $k \geq 1$ and show the validity of (3.1c) for $k+1$. Because of $C_{k+1} = B_{k+1} \cap C_k$ and Proposition 1.5(ii),

$$\lambda_c(B_{k+1} \cup C_k) = \lambda_c(B_{k+1}) + \lambda_c(C_k) - \lambda_c(C_{k+1}). \qquad (3.1e)$$

Due to (3.1e), the inequality $\lambda_c(B_{k+1}) \geq \lambda_c(A_{k+1}) - \frac{1}{2^{k+1}}\varepsilon$ [by (3.1b) for $k+1$], and the assumption about the validity of (3.1c) for some k, we have:

$$\lambda_c(B_{k+1} \cup C_k) \geq \lambda_c(A_{k+1})$$

$$+ \lambda_c(A_k) - \lambda_c(C_{k+1}) - \varepsilon\big(1 - \tfrac{1}{2^{k+1}}\big).$$

Since obviously $B_{k+1} \cup C_k \subseteq A_k$, we have

$$\lambda_c(A_k) \geq \lambda_c(B_{k+1} \cup C_k),$$

and hence

$$\lambda_c(C_{k+1}) \geq \lambda_c(A_{k+1}) - \varepsilon\big(1 - \tfrac{1}{2^{k+1}}\big)$$

$$+ \lambda_c(A_k) - \lambda_c(B_{k+1} \cup C_k)$$

$$\geq \lambda_c(A_{k+1}) - \varepsilon\big(1 - \tfrac{1}{2^{k+1}}\big).$$

This proves (3.1c) and (3.1d) and thereby yields that $\{\overline{C}_k\}$ is a monotone nonincreasing sequence of nonempty compact sets; hence, by Cantor's Theorem 5.4, Chapter 2,

$$\bigcap_{k=1}^{\infty} A_k \supseteq \bigcap_{k=1}^{\infty} \overline{C}_k \neq \emptyset.$$

In conclusion, λ_c is indeed a premeasure on the ring $\mathcal{R}$. □

3.2 Remarks and Definitions.

(i) Theorem 3.1 states that the Lebesgue content on $\mathcal{R}(\mathcal{S})$ in $\mathbb{R}^n$ is σ-additive. This, obviously, implies that the Lebesgue elementary content is also σ-additive on $\mathcal{S}$.

(ii) In Example $1.2(ii)$ we defined the Lebesgue elementary content λ^0 on the semi-ring $\mathcal{S}$ of half-open intervals in $\mathbb{R}^n$. Now, by the use of Proposition 2.11, Corollary 2.14, and Theorem 3.1, we can have the couple $(\mathcal{R}, \lambda_c)$ or, in light of Remark (i), even $(\mathcal{S}, \lambda^0)$ as an extendible formatter of the outer measure λ^* acting on $\mathcal{P}(\mathbb{R}^n)$. We call this set function the *Lebesgue outer measure*. The σ-algebra $\Sigma^* \subseteq \mathcal{P}(\mathbb{R}^n)$ of all λ^*-measurable sets, in the specifically designated notation $\mathcal{L}^*$, is called the *Lebesgue σ-algebra of measurable sets.* The corresponding restriction of λ^* to $\mathcal{L}^*$, $\lambda_0^* = \mathrm{Res}_{\mathcal{L}^*}\lambda^*$, is called the *Lebesgue measure.* The *Lebesgue measure space* $(\mathbb{R}^n, \mathcal{L}^*, \lambda_0^*)$ will form a complete measure space, as per Proposition 2.7. The further restriction $\lambda = \mathrm{Res}_{\Sigma(\mathcal{S})}\lambda^* = \mathrm{Res}_{\mathcal{B}(\mathbb{R}^n)}\lambda^*$ of the Lebesgue outer measure to the smallest σ-algebra generated by $\mathcal{S}$ (which, according to Theorem 2.8, Chapter 4, is identical to the smallest σ-algebra generated by the usual topology) or, equivalently, by $\mathcal{R}$, known as the Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}^n$, is referred to as the *Borel - Lebesgue measure*. The reader should be aware of the weighty differences between these two measures and measure spaces, often confused with one another in the literature on measure theory.

By noticing that there exists a monotone increasing sequence $(-k, k]^n \uparrow \mathbb{R}^n$ of half-open squares with

$$\lambda^0((-k, k]^n) = 2^n k^n < \infty$$

we conclude that λ^0 is σ-finite on $\mathcal{S}$ and, therefore, by Corollary 2.14, the Borel - Lebesgue measure λ is unique on $\mathcal{B}$. By Remark 2.24, the Lebesgue outer measure λ^* and hence the Lebesgue measure λ_0^* are also unique on $\mathcal{P}(\mathbb{R}^n)$ and $\mathcal{L}^*$, respectively. Finally, by Theorem 2.19 (ii), the completion of Borel - Lebesgue measure λ coincides with Lebesgue measure λ_0^* on $\mathcal{L}^*$ and the corresponding completion of the Borel σ-algebra coincides with the σ-algebra $\mathcal{L}^*$ of Lebesgue measurable sets.

Both Lebesgue and Borel - Lebesgue measures have their strengths and weaknesses. The Borel - Lebesgue measure operates on the Borel σ-algebra, which stems from the usual topology and preserves reasonable topological properties. Furthermore, the Borel - Lebesgue measure is an element of a very important class of Borel measures. However, unlike the Lebesgue measure, the Borel - Lebesgue measure is not complete. □

3.3 Definitions.

(i) Let $\mathcal{B} = \Sigma(\tau)$ be a Borel σ-algebra in Ω. Any measure μ on $\mathcal{B}$ is called a *Borel measure* and the triple $(\Omega, \mathcal{B}, \mu)$ is called a *Borel measure space*.

(ii) Let (Ω, d) be a metric space and $\mathcal{B}$ be the Borel σ-algebra induced by the metric topology. A Borel measure μ on $\mathcal{B}$ is said to be a *Borel - Stieltjes measure* if $\mu(B) < \infty$ for any d-bounded Borel set B. (More on Borel - Stieltjes measures in Problem 3.17.)

A Borel measure μ on $(\mathbb{R}^n, \mathcal{B})$ is said to be a *Borel - Lebesgue - Stieltjes measure* if $\mu(B) < \infty$ for any d_e-bounded Borel set B. Clearly, any Borel - Lebesgue - Stieltjes measure is σ-finite. In other words, a Borel - Lebesgue - Stieltjes measure is a Borel - Stieltjes measure on $\mathcal{B}$ that is generated by the Euclidean (usual) metric topology.

(iii) Let μ be a Borel - Lebesgue - Stieltjes measure on $(\mathbb{R}^n, \mathcal{B})$. Now, in light of Carathéodory's construction we can use the couple $(\mathcal{B}, \mu)$ as an extendible formatter of the outer measure μ^* operating on $\mathcal{P}(\mathbb{R}^n)$ and call this set function the *Borel outer measure*. The σ-algebra $\Sigma^* \subseteq \mathcal{P}(\mathbb{R}^n)$ of all μ^*-measurable sets is denoted by $\mathcal{B}_\mu^*$ and called the *Lebesgue - Stieltjes σ-algebra of measurable sets*. The corresponding restriction $\mu_0^* = \text{Res}_{\mathcal{B}_\mu^*}\mu^*$ is called the *Lebesgue - Stieltjes measure*. In the literature on measure theory, the Lebesgue - Stieltjes measure is often confused with the Borel - Lebesgue - Stieltjes measure. □

In addition to the Borel - Lebesgue measure on Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}^n$, we present another construction of a Borel - Lebesgue - Stieltjes measure, for simplicity letting dimension $n = 1$. In Example

1.2(iii), we introduced the Lebesgue - Stieltjes elementary content μ_f^0 on the semi-ring $\mathcal{S}$ of all half-open intervals $(a, b] \subseteq \mathbb{R}$, by means of an extended distribution function (i.e., a monotone nondecreasing, right-continuous function) $f\colon\ (\mathbb{R}, d_e) \to (\mathbb{R}, d_e)$, as $\mu_f^0((a, b]) = f(b) - f(a)$. Obviously, μ_f^0 reduces to the Lebesgue elementary content if $f(x) = x$. According to Proposition 2.11, μ_f^0 can uniquely be extended to the *Lebesgue - Stieltjes content* μ_f on the ring $\mathcal{R}(\mathcal{S})$ of "figures." The following is to show that μ_f is σ-additive.

3.4 Theorem. *Let μ_f be a Lebesgue - Stieltjes content on the ring $\mathcal{R}(\mathcal{S})$ induced by a monotone nondecreasing right-continuous function f. Then μ_f is a premeasure.*

Proof. Because μ_f is finite on $\mathcal{R}(\mathcal{S})$, as in Theorem 3.1, it is sufficient to show that μ_f is Ø-continuous. Let $\{R_n\}$ be a sequence of sets from $\mathcal{R}(\mathcal{S})$ monotonically decreasing to Ø. We prove that $\lim_{n \to \infty} \mu_f(R_n) = 0$.

We assume that $R_n \subseteq C,\ n = 1, 2, \dots,$ where C is a compact set in $(\mathbb{R}, \tau_e)$. A set $R_n \in \mathcal{R}$ is a *figure* if it is a finite union of disjoint intervals of type $(a, b]$. Because of right-continuity of f, it can be easily shown that, for each fixed $\varepsilon > 0$ and for any figure R_n, there is a subfigure $B_n \subseteq R_n$ such that $\overline{B}_n \subseteq R_n$ and such that $\mu_f(R_n) - \mu_f(B_n) < \varepsilon 2^{-n}$. It also follows that $\bigcap_{n=1}^{\infty} \overline{B}_n = \text{Ø}$. We claim that there is an r such that $\bigcap_{k=1}^{r} \overline{B}_k = \text{Ø}$. To see this, observe that $\{C \backslash \overline{B}_n = C \cap (\overline{B}_n)^c;\ n = 1, 2, \dots\}$ is an open cover of C in the relative topology $(C,\ \tau_e \cap C)$. Because compactness is weakly hereditary and C is closed, it follows that C is also compact in $\tau_e \cap C$. Thus, the above cover reduces to a finite subcover, say $C \backslash \overline{B}_1, \dots, C \backslash \overline{B}_r$, yielding that

$$C = \bigcup_{k=1}^{r} (C \backslash \overline{B}_k) = C \backslash (\bigcap_{k=1}^{r} \overline{B}_k).$$

Thus,

$$\bigcap_{k=1}^{r} \overline{B}_k = \emptyset \quad \text{and} \quad \bigcap_{k=1}^{r} B_k = \emptyset.$$

Now, for all $n \geq r$,

$$\bigcap_{k=1}^{n} B_k = \emptyset$$

and

$$\mu_f(R_n) = \mu_f(R_n \setminus \bigcap_{k=1}^{n} B_k) + \mu_f(\bigcap_{k=1}^{n} B_k) = \mu_f(R_n \setminus \bigcap_{k=1}^{n} B_k).$$

In as much as $\{R_n\}$ is monotone decreasing, it follows that

$$R_n \setminus (\bigcap_{k=1}^{n} B_k) = \bigcup_{k=1}^{n} (R_n \setminus B_k) \subseteq \bigcup_{k=1}^{n} (R_k \setminus B_k).$$

Observe that this is the desired inclusion implying the estimate $\mu_f(R_n) < \varepsilon$. This inclusion is due to the inclusion $R_n \setminus B_k \subseteq R_k \setminus B_k$, which holds for all $k \leq n$ (as long as $n < \infty$). Hence, the above countable intersection reduces to a finite intersection of the sets $\overline{B}_k,\ k = 1, \ldots, r$. Thus we have

$$\mu_f(R_n) = \mu_f(R_n \setminus (\bigcap_{k=1}^{n} B_k))$$

$$\leq \mu_f(\bigcup_{k=1}^{n} R_k \setminus B_k) \leq \sum_{k=1}^{n} \mu_f(R_k \setminus B_k) \leq \varepsilon(1 - 2^{-n}),$$

which implies that $\mu_f(R_n) \to 0$. □

It can alternatively be shown that the Lebesgue - Stieltjes elementary content is σ-additive with no prior extension to the Lebesgue - Stieltjes content and bypassing Ø-continuity.

3.5 Remarks.

(i) Using the same arguments as in Remarks and Definitions 3.2, we will extend the Lebesgue - Stieltjes elementary content μ_f^0 (or content μ_f) from the semi-ring $\mathcal{S}$ (or ring $\mathcal{R}(\mathcal{S})$, respectively) to the *Lebesgue - Stieltjes measure* μ_f^* on the σ-algebra Σ^* $(= \mathcal{B}_\mu^*)$ of *Lebesgue - Stieltjes measurable sets* and then reduce it to the unique

measure μ_f, which is clearly a Borel - Lebesgue - Stieltjes measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$.

(ii) When dealing with Borel measures, it is common to observe a certain property of a σ-finite Borel measure μ_1 on the semi-ring $\mathcal{S}$ in $\mathbb{R}^n$ and extend this property of μ_1 from $\mathcal{S}$ to the Borel σ-algebra $\mathcal{B}$ arriving at another Borel measure μ_2. Because μ_1 and μ_2 coincide on $\mathcal{S}$, by Corollary 2.14, $\mu_1 = \mu_2$ on $\mathcal{B}$. Consequently, by Remark 2.24, the corresponding outer measures μ_1^* and μ_2^* must coincide on $\mathcal{P}(\mathbb{R}^n)$ as well as their restrictions on $\mathcal{B}_1^* = \mathcal{B}_2^*$. Note, however, that $\mathcal{B}_\mu^*$ is not a general notation as $\mathcal{B}$ is, for it is not induced by the usual topology and it is related to a particular Borel measure μ on $\mathcal{B}$.

(iii) We have learned that if f is an extended distribution function [see the definition in Example 1.2(iii)], then it induces a Borel - Lebesgue - Stieltjes measure on $\mathcal{B}$. Conversely, a Borel - Lebesgue - Stieltjes measure μ generates an extended distribution function.

If μ is a finite Borel - Lebesgue - Stieltjes measure on $\mathcal{B}$, then we can set $f(x) = \mu((-\infty, x])$ and such an f is a distribution function. Indeed, take a sequence $x_1 > x_2 > \cdots \to x$. Then $f(x_n) - f(x) = \mu((x, x_n]) \to 0$, by Ø-continuity of μ [Theorem 1.7 (i)], which shows that f is right-continuous. Since μ is a finite measure, f is bounded. Finally, if x_n is any monotone decreasing sequence convergent to $-\infty$ (such as $\{-n\}$), then, again by Ø-continuity of μ, it follows that $\mu((-\infty, x_n])$ and thus, $f(x_n) \to 0$.

If μ is an arbitrary Borel - Lebesgue - Stieltjes measure, we can define $f(0) = 0$ and

$$f(x) = \begin{cases} \mu((0, x]), & x > 0 \\ -\mu((x, 0]), & x < 0. \end{cases} \tag{3.5}$$

Similarly, one can show that f is an extended distribution function. (See Problem 3.3.) If $\mathfrak{B} = \mathfrak{B}(\mathbb{R}, \mathcal{B})$ denotes the set of all Borel - Lebesgue - Stieltjes measures on $(\mathbb{R}, \mathcal{B})$, then it can be shown that any two extended distribution functions f_1 and f_2 that induce $\mu \in \mathfrak{B}$ can differ only in an additive constant (see Problem 3.4). The latter

generates an equivalence relation among all extended distribution functions, say E. So, the distribution functions that differ only in an additive constant form a class associated with one such measure μ. Therefore, if $\mathfrak{D}_e$ denotes the set of all extended distribution functions, for each $\mu \in \mathfrak{B}$, there is a unique equivalence class $\{f; \mu\}$ of all extended distribution functions that induce μ, such that $\{f; \mu\} = \{f + c : c \in \mathbb{R}\}$. Let $\mathfrak{D}_e|\mathrm{E} := \{\{f; \mu\} : \mu \in \mathfrak{B}\}$ be the corresponding quotient set of $\mathfrak{D}_e$. Consequently, there is a bijective map $\mathcal{D}$ from set $\mathfrak{B}$ onto set $\mathfrak{D}_e|\mathrm{E}$.

As regards the subset $\mathfrak{B}_*$ of all finite Borel - Lebesgue - Stieltjes measures, then, obviously, each one of them generates a unique distribution function and there is a bijective map between $\mathfrak{B}_*$ and the set $\mathfrak{D}$ ($\subseteq \mathfrak{D}_e$) of all distribution functions.

To make all distinctions between distribution and extended distribution functions lucid the reader may find it plausible to go over Problem 3.9. □

We return to the Lebesgue measure λ_0^* on $\mathcal{L}^*$. First, we prove a lemma about negligible sets. One of the interesting consequences of this result is that in $\mathbb{R}^n$, all Borel sets having a dimension less than n are null sets.

3.6 Lemma. *A set $N \subseteq \mathbb{R}^n$ is λ-negligible if and only if for each $\varepsilon > 0$ there is a countable cover $\{I_k\} \subseteq \mathcal{S}$ of N by semi-open intervals I_k's such that $\sum_{k=1}^{\infty} \lambda^0(I_k) < \varepsilon$.*

Proof. Let N be λ-negligible. Then, by Problem 2.18, $\lambda^*(N) = 0$ and

$$0 = \lambda^*(N) = \inf\{\textstyle\sum_{k=1}^{\infty} \lambda^0(I_k) \colon \{I_k\} \in \mathscr{C}_N\},$$

where $\mathscr{C}_N$ is the set of all countable covers of N by semi-open intervals and it is nonempty, because otherwise $\lambda^*(N)$ would equal ∞. By the definition of a limit inferior, for each $\varepsilon > 0$, there is a cover $\{I_k\} \in \mathscr{C}_N$ such that

$$\lambda^*(N) + \varepsilon > \textstyle\sum_{k=1}^{\infty} \lambda^0(I_k),$$

which proves the first part of the statement.

Conversely, let $\varepsilon > 0$ and let $\{I_k\} \subseteq \mathcal{S}$ be a countable cover of N with the property that $\sum_{k=1}^{\infty} \lambda^0(I_k) < \varepsilon$. Then,

$$\lambda^*(N) \leq \lambda^*(\bigcup_{k=1}^{\infty} I_k) \leq \sum_{k=1}^{\infty} \lambda^0(I_k) < \varepsilon$$

and hence, by Problem 2.18, N is a λ-negligible set. □

3.7 Lemma. *Let $f: \mathbb{R} \to \mathbb{R}$ be an additive function, continuous at zero. Then, f is linear.*

Proof. First note that

$$f(0) + f(0) = f(0 + 0) = f(0). \tag{3.7a}$$

This yields that $f(0) = 0$. Then, from

$$0 = f(0) = f(x - x) = f(x) + f(-x) \tag{3.7b}$$

it follows that $f(x) = -f(-x)$ and thus f is odd. Now, let n be any positive integer number. Then, in as much as f is additive,

$$f(nx) = nf(x). \tag{3.7c}$$

If n is a negative integer, then, from (3.7b - c),

$$\begin{aligned} f(nx) &= f((-n)(-1)x) \\ &= -f(-nx) = -(-n)f(x) = nf(x). \end{aligned}$$

Hence, for each $n \in \mathbb{Z}$,

$$f(nx) = nf(x), \tag{3.7d}$$

which yields that

$$f(\tfrac{x}{n}) = \tfrac{1}{n} f(x). \tag{3.7f}$$

Combining (3.7d) and (3.7f) we have that for each integer m,

$$mf(\tfrac{x}{n}) = f(\tfrac{m}{n}x) = \tfrac{m}{n}f(x).$$

In other words, for each rational number q,

$$f(qx) = qf(x).$$

Because f is continuous at zero and because f is additive and odd we have from

$$f(x-y) = f(x) + f(-y) = f(x) - f(y)$$

that f is continuous on $\mathbb{R}$. Now, let $r \in \mathbb{R}$. Then, there is a sequence $\{q_n\}$ of rationals convergent to r. Due to continuity of f,

$$\lim_{n\to\infty} f(q_n) = f(r). \tag{3.7g}$$

On the other hand, $f(q_n \cdot 1) = q_n f(1)$ and (3.7g) lead to

$$f(r) = \lim_{n\to\infty} f(q_n) = f(1) \lim_{n\to\infty} q_n = f(1)r.$$

This shows that f is a linear function $f(x) = cx,$ where $c = f(1)$. □

3.8 Corollary. *Let $f: \mathbb{R}^n \to \mathbb{R}$ be continuous at zero and additive for each variable separately. Then*

$$f(x_1, \ldots, x_n) = cx_1 \cdots x_n, \text{ where } c = f(1, \ldots, 1).$$

Proof. If $x_2, \ldots, x_n$ are fixed, then by Lemma 3.7,

$$f(x_1, \ldots, x_n) = x_1 f(1, x_2, \ldots, x_n).$$

Applying the same procedure successively to the other variables we have the statement. □

3.9 Definition. Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a bijective function such that $L(y) = y + x,$ for a fixed $x \in \mathbb{R}^n$. A Borel measure μ on $\mathcal{B}(\mathbb{R}^n)$ is said to be *translation-invariant*, if for each Borel set $B \in \mathcal{B}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, $\mu(L_*(B)) = \mu(B)$, where $L_*(B) = B + x$

$= \{x + y : y \in B\}$. Now, $L_*(B) = B + x$ is Borel, because L is an affine (that is, continuous) transformation. □

We show in Section 4 that the Borel - Lebesgue and Lebesgue measures are translation-invariant. The following theorem states that any translation-invariant Borel measure is a multiple of the Borel - Lebesgue measure.

3.10 Theorem. *Let μ be a translation-invariant Borel measure on $\mathcal{B}(\mathbb{R}^n)$. Then, $\mu = c\lambda$, where λ is the* Borel - Lebesgue *measure on $\mathcal{B}(\mathbb{R}^n)$ and $c = \mu(C)$ (C stands for a unit cube*).

Proof. For each $x \in \mathbb{R}$, define

$$I_x = \begin{cases} [x, 0), & x < 0 \\ \emptyset, & x = 0 \\ [0, x), & x > 0 \end{cases}$$

and

$$\operatorname{sgn} x = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0. \end{cases}$$

Denote

$$f(x_1, \ldots, x_n) = \operatorname{sgn}\left(\prod_{i=1}^{n} x_i\right) \mu\left(\prod_{i=1}^{n} I_{x_i}\right). \tag{3.10}$$

We show that f defined in (3.10) is additive and continuous in each variable separately. Without loss of generality, we show it with respect to x_1. Let $x_1 = x + y$.

Case 1. Suppose $x > 0$ and $y > 0$. Then,

$$I_{x+y} = [0, x + y) = [0, x) + [x, x + y)$$

and

$$\prod_{i=1}^{n} I_{x_i} = R_1 + R_2,$$

where

$$R_1 = I_x \times I_{x_2} \times \cdots \times I_{x_n}$$

and

$$R_2 = [x, x+y) \times I_{x_2} \times \cdots \times I_{x_n}.$$

In as much as $x, y,$ and $x+y$ are all positive,

$$\operatorname{sgn}\left(\prod_{i=1}^{n} x_i\right) = \operatorname{sgn}(x \cdot x_2 \cdots x_n)$$

$$= \operatorname{sgn}(y \cdot x_2 \cdots x_n). \tag{3.10a}$$

From (3.10a),

$$f(x_1, \ldots, x_n) = f(x+y, x_2, \ldots, x_n)$$

$$= \operatorname{sgn}(x \cdot x_2 \cdots x_n)\mu(R_1) = \operatorname{sgn}(y \cdot x_2 \cdots x_n)\mu(R_2)$$

and because μ is translation-invariant,

$$f(x_1, \ldots, x_n) = f(x, x_2, \ldots, x_n) + f(y, x_2, \ldots, x_n).$$

Case 2. Suppose $x+y > 0$ and $x > 0,\ y < 0$. Then,

$$\operatorname{sgn}((x+y) \cdot x_2 \cdots x_n)$$

$$= \operatorname{sgn}(x \cdot x_2 \cdots x_n) = -\operatorname{sgn}(y \cdot x_2 \cdots x_n). \tag{3.10b}$$

Because

$$I_{x+y} = [0, x+y) = [0, x) \backslash [x+y, x),$$

we have

$$\lambda([x+y, x)) = \lambda([y, 0)),$$

and because μ is translation-invariant, using (3.10b) we have that

$$
\begin{aligned}
& f(x+y, x_2, \ldots, x_n) \\
&= \operatorname{sgn}((x+y)\cdot x_2 \cdots x_n) \\
&\quad \mu(I_x \times I_{x_2} \times \cdots \times I_{x_n}) - \mu(I_y \times I_{x_2} \times \cdots \times I_{x_n}) \\
&= \operatorname{sgn}(x \cdot x_2 \cdots x_n)\mu(I_x \times I_{x_2} \times \cdots \times I_{x_n}) \\
&+ \operatorname{sgn}(y \cdot x_2 \cdots x_n)\mu(I_y \times I_{x_2} \times \cdots \times I_{x_n}) \\
&= f(x, x_2, \ldots, x_n) + f(y, x_2, \ldots, x_n).
\end{aligned}
$$

Case 3. $x + y > 0$ and $x < 0,\ y > 0$ are the same as in case 2.

The other combinations of x and y are left for the reader. (See Problem 3.23.)

Now, we prove continuity of f at zero. Let $\{a_k\}$ be a sequence convergent to zero from the right. Then, $\{a_k\} \subseteq \mathbb{R}_+$ and the sequence of sets $\{I_{a_k}\}$ is such that $\bigcap_{k=1}^{\infty} I_{a_k} = \{0\}$. The latter yields that

$$
\bigcap_{k=1}^{\infty} \{I_{a_k} \times I_{x_2} \times \cdots \times I_{x_n}\} = I_0 \times I_{x_2} \times \cdots \times I_{x_n}.
$$

By the definition, $I_0 = \emptyset$; and by continuity from above of μ, we have that

$$
\lim_{k \to \infty} f(a_k, x_2, \ldots, x_n) = 0.
$$

Similarly, by continuity from below of μ, we have that

$$
\lim_{k \to \infty} f(a_k, x_2, \ldots, x_n) = 0
$$

for $\{a_k\} \uparrow 0$. In addition, $f(0, x_2, \ldots, x_n) = 0$ holding true by the definition of f.

By Corollary 3.8,

$$f(x_1, \ldots, x_n) = f(1, \ldots, 1)x_1 \cdots x_n$$

$$= \underbrace{\text{sgn}(1 \cdots 1)}_{1} x_1 \cdots x_n \mu(C), \tag{3.10c}$$

where $C = [0, 1) \times \cdots \times [0, 1)$. On the other hand,

$$f(x_1, \ldots, x_n) = \text{sgn} \left(\prod_{i=1}^{n} x_i \right) \mu \left(\prod_{i=1}^{n} I_{x_i} \right),$$

which, along with (3.10c), gives

$$\mu \left(\prod_{i=1}^{n} I_{x_i} \right) = \frac{x_1 \cdots x_n}{\text{sgn}(x_1 \cdots x_n)} \mu(C). \tag{3.10d}$$

Note that

$$\lambda^0 \left(\prod_{i=1}^{n} I_{x_i} \right) = \frac{x_1 \cdots x_n}{\text{sgn}(x_1 \cdots x_n)}. \tag{3.10e}$$

Equations (3.10d) and (3.10e) tell us that for any rectangle R whose all sides lie on corresponding coordinate axes,

$$\mu(R) = \mu(C)\lambda^0(R). \tag{3.10f}$$

For an arbitrarily positioned rectangle R all of whose sides are parallel to the corresponding coordinate axes, (3.10f) still holds true due to the translation-invariance of μ.

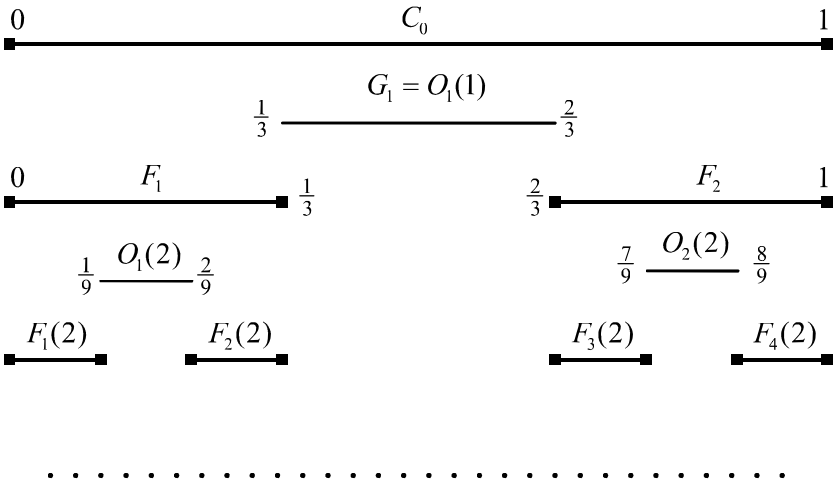
By $\mu_0 = \mu(C)\lambda^0$ we define an elementary content on the semi-ring $\mathcal{S}$ of half-open rectangles. Then, by $\widetilde{\mu} = \mu(C)\lambda$ we also have a Borel measure on $\mathcal{B}$. Now, we have three Borel measures on $\mathcal{B}$: $\widetilde{\mu}$, μ, and the (unique, as $\mathcal{S}$ is $\cap$-stable) extension of μ_0 from $\mathcal{S}$ to $\mathcal{B}$. All three coincide on $\mathcal{S}$ and therefore must be equal on $\mathcal{B}$. □

3.11 Example. (Cantor ternary set). Consider the following family of subsets of $[0, 1]$. Let

$$\Omega = C_0 = [0,1],\ G_1 = (\tfrac{1}{3}\,,\ \tfrac{2}{3}),\ C_1 = C_0 \backslash G_1,$$

$$G_2 = (\tfrac{1}{9}\,,\ \tfrac{2}{9}) \cup (\tfrac{7}{9}\,,\ \tfrac{8}{9}),\ C_2 = C_1 \backslash G_2,$$

as depicted in Figure 3.1 below:



. . . . . .

Figure 3.1

Therefore, each C_n is the union of 2^n closed intervals, and each G_n is the union of 2^{n-1} open intervals. Also,

$$C_{n+1} = C_n \backslash G_{n+1} = C_n \backslash \bigcup_{k=1}^{2^n} O_k(n+1) = \bigcup_{k=1}^{2^{n+1}} F_k(n)$$

and $\{C_n\}$ is a monotone decreasing sequence of sets. The *Cantor set* is defined as

$$C = \bigcap_{n=1}^{\infty} C_n$$

and it can be characterized as follows.

1) C is closed as the intersection of closed sets.

2) Each C_n contains 2^n closed disjoint intervals $F_1(n), \ldots,$ $F_{2^n}(n)$. Each of these intervals is a term of the monotone decreasing sequence $\{F_k(n)\}\downarrow$ with $d_e(F_k(n)) = \lambda(F_k(n))\downarrow 0,\ n \to \infty$. By applying Cantor's Theorem 5.4, Chapter 2, we conclude that $\forall$ $k = 1, 2, \ldots, \bigcap_{n=1}^{\infty} F_k(n)$ consists of exactly one point. In other words, C is a union of isolated points and therefore nowhere dense.

3) $C_n = C_0 \setminus \bigcup_{k=1}^{n} G_k = \left(\bigcup_{k=1}^{n} G_k\right)$

4) $G = C^c = \bigcup_{n=1}^{\infty} C_n^c = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{n} G_k = \sum_{n=1}^{\infty} G_n$.

5) The Lebesgue measure of G_n is $\lambda(G_n) = \frac{1}{2}(\frac{2}{3})^n$, because

$$\lambda(G_1) = \tfrac{1}{3}\,,\ \lambda(G_2) = 2\tfrac{1}{3^2}\,,\ \lambda(G_3) = 4\tfrac{1}{3^3} = \tfrac{1}{2}(\tfrac{2}{3})^3, \ldots\ .$$

6) Thus,

$$\lambda(G) = \textstyle\sum_{n=1}^{\infty}\lambda(G_n) = \frac{1}{3}\sum_{n=0}^{\infty}(\frac{2}{3})^n = 1.$$

Hence $\lambda(C) = 1 - 1 = 0$ and therefore C is a Borel λ-null set.

7) C is not empty, because C contains all boundary points of the sets C_n: $0, 1, \frac{1}{3}, \frac{2}{3}, \frac{1}{3^2}, \frac{2}{3^2}, \frac{7}{3^2}, \frac{8}{3^2}, \ldots$.

These boundary points have the following ternary representations:

$0 = 0.0000\ldots$

$1 = 1.0\ (\text{or})\ = \ 0.22222\ldots\ (\text{in dyadic representation})$

$\frac{1}{3} = 0.1\ (\text{or})\ = 0.02222\ldots$

$\frac{2}{3} = 0.2 = 0.1222\ldots$

$\frac{1}{9} = 0.01 = 0.00222\ldots$

$\frac{2}{9} = 0.02 = 0.0122\ldots$

$\frac{7}{9} = 0.21 = 0.2022\ldots$

. . . . . .

Every set C_n has exactly 2^n boundary points, each one of which has a unique triadic representation consisting of all n-tuples of digits 0 or 2. Observe that $2^n = |\mathcal{P}(A)|$ where A is an n-element set. Therefore, C is equivalent to the set of all subsets of natural numbers which, by Remark 6.5 (i), has the cardinality of the continuum. In other words, $C \approx \mathbb{R}$. Therefore, the Cantor set is an example of a uncountable Borel λ-null set. □

The following regularity notions are used in several problems below and will be revisited in [Dsh2] for more general measures on topological spaces (with a whole chapter dedicated to this theme).

3.12 Definitions. Let μ be a Borel measure on a Borel measurable space $(\Omega, \mathcal{B}(\tau))$ such that τ is a Hausdorff topology on Ω.

(i) A Borel set $A \in \mathcal{B} = \mathcal{B}(\tau)$ is called

- (a) *μ-outer regular* if $\mu(A) = \inf\{\mu(O) : A \subseteq O \in \tau\}$,
- (b) *μ-inner regular* if $\mu(A) = \sup\{\mu(K) : A \supseteq K \in \mathcal{K}\}$ where $\mathcal{K}(\Omega)$ is the family of all compact subsets of Ω,
- (c) *μ-closed regular* if $\mu(A) = \sup\{\mu(F) : A \supseteq F \in \tau^c\}$.

(ii) A Borel measure μ is *outer* (*inner* or *closed*) *regular* on $\mathcal{G} \subseteq \mathcal{B}(\Omega)$ if all sets of $\mathcal{G}$ are μ-outer (μ-inner or μ-closed) regular.

(iii) μ is *Radon* if μ is

- (a) finite on $\mathcal{K}(\Omega)$;
- (b) outer regular on $\mathcal{B}(\Omega)$;
- (c) inner regular on τ.

(iv) μ is *regular* (*weakly regular*) if it is outer and inner (closed) regular on $\mathcal{B}(\Omega)$.

(v) μ is *strongly regular* if it is regular and finite on $\mathcal{K}(\Omega)$. □

PROBLEMS

3.1 Let $H = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n\colon\ x_i = a \in \mathbb{R}\}$ be a hyperplane orthogonal to the ith coordinate axis. Show that H is a λ-null Borel set. [*Hint*: 1) Show that H is closed in $(\mathbb{R}^n, \tau_e)$ and hence Borel. 2) Find a relevant countable cover of H by rectangles from $\mathcal{S}$ and apply Lemma 3.6.]

3.2 Show that each countable subset of $\mathbb{R}^n$ is a Borel λ^n-null set.

3.3 Show that f defined by (3.5) in Remark 3.5(iii) is an extended distribution function.

3.4 Let f_1 and f_2 be two extended distribution functions and let μ_1 and μ_2 be the corresponding Borel - Lebesgue - Stieltjes measures induced by these functions. Show that $\mu_1 = \mu_2$ if and only if $f_1 - f_2 = \boldsymbol{c},$ where $\boldsymbol{c}$ is a constant function.

3.5 Let $\mathfrak{D}_e$ be the set of all extended distribution functions. Show that $\mathfrak{D}_e$ is a semivector space over $\mathbb{R}_+$.

3.6 Let f_1 and f_2 be two extended distribution functions. If μ_1 and μ_2 are the corresponding Borel - Lebesgue - Stieltjes measures induced by f_1 and $f_2,$ show that for any nonnegative scalars α_1 and α_2, $\mathcal{D}(\alpha_1\mu_1 + \alpha_2\mu_2) = \{\alpha_1 f_1 + \alpha_2 f_2; \mu\}$, where $\mathcal{D}$ is defined in Remark 3.5(iii) and $\mu = \alpha_1\mu_1 + \alpha_2\mu_2$.

3.7 Let $f\colon \mathbb{R} \to \mathbb{R}$ be an extended distribution function and let μ_f be the corresponding Borel - Lebesgue - Stieltjes measure on $\mathcal{B}(\mathbb{R})$. Show that

$a)$ $\mu_f((a, b)) = f(b-) - f(a)$.

$b)$ $\mu_f([a, b]) = f(b) - f(a-)$.

$c)$ $\mu_f([a, b)) = f(b-) - f(a-)$.

$d)$ f is continuous if and only if $\mu_f(\{x\}) = 0,\ x \in \mathbb{R}$.

3.8 Let f be the extended distribution function on $\mathbb{R}$ given by

$$f(x) = \begin{cases} -1, & x < -2 \\ 1+x, & -2 \le x < 0 \\ 2+x^2, & 0 \le x < 3 \\ 15, & 3 \le x \end{cases}$$

and let μ_f be the corresponding Borel - Lebesgue - Stieltjes measure. Evaluate the measure of the following sets.

$a)\ \{3\} \quad b)\ \{-2\} \quad c)\ [-1, 3) \quad d)\ [0, \frac{1}{2}) \cup (3, 5).$

3.9 Let f be a distribution function and let μ_f denote the Borel - Lebesgue - Stieltjes measure induced by f. Justify with a proof or give a counterargument:

$a)$ Must f be an extended distribution function?

$b)$ Suppose g is a function defined by (3.5) of Remark 3.5 (iii). Is g a distribution function? If your answer is yes, is $g = f$?

3.10 Let μ be an atomic measure ($= \sum_{i=0}^{\infty} \alpha_i \varepsilon_{b_i}$).

1) Is μ always a Borel - Lebesgue - Stieltjes measure? If it is not, give a condition under which μ is a Borel - Lebesgue - Stieltjes measure.

2) In this case, give $\{f; \mu\}$.

3) Plot one such f.

3.11 Consider the Borel σ-algebra $\mathcal{B} = \mathcal{B}(\mathbb{R}^n)$ generated by the usual topology. Show that, for any Borel set $B \in \mathcal{B}$ and any point $x \in \mathbb{R}^n$, $B + x = \{z \in \mathbb{R}^n \colon z = y + x \colon y \in B\} \in \mathcal{B}$.

3.12 Let $(\Omega, \mathcal{B}, \mu)$ be a Borel measure space, such that the Borel σ-algebra $\mathcal{B}$ is generated by a Hausdorff topological space τ, and μ is a finite Borel measure on $\mathcal{B}$. For any subset $Q \subseteq \Omega$, denote by $\mathcal{K}(Q)$ the collection of all compact subsets of Q. Show that the family $\mathcal{M} \subseteq \mathcal{B}$ of all inner regular sets $B \in \mathcal{B}$, i.e., with

$$\mu(B) = \sup\{\mu(K) \colon K \in \mathcal{K}(B)\},$$

forms a monotone system in Ω.

3.13 Let $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ be a finite Borel measure space (i.e., μ is finite). Show that μ is weakly regular, that is, it is outer and closed regular on $\mathcal{B}(\mathbb{R}^n)$.

3.14 Let $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ be a finite Borel measure space. Show that μ is strongly regular.

3.15 Suppose that $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ is a σ-finite Borel measure space. Show μ is inner regular.

3.16 Let $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ be a finite Borel measure space. Use just the inner regularity of μ to prove that μ is outer regular.

3.17 Let (Ω, d) be a metric space and $\mathcal{B}$ be the Borel σ-algebra generated by the metric topology. Furthermore, let μ be a Borel - Stieltjes measure on $\mathcal{B}$. Is μ σ-finite?

3.18 Show that a σ-finite Borel measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, need not be outer regular.

3.19 Let $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$ be the Borel - Lebesgue measure space. Show that the Borel - Lebesgue measure λ is strongly regular.

3.20 Let λ denote the Borel - Lebesgue measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$. Show that for each Borel set B and $\varepsilon > 0$, there is a countable cover of B by disjoint half-open cubes $\{C_k\}$ such that

$$\sum_{k=1}^{\infty} \lambda^0(C_k) - \lambda(B) < \varepsilon.$$

In particular,

$$\lambda(B) = \inf\{\textstyle\sum_{k=1}^{\infty} \lambda^0(C_k) \colon \{C_k\} \in \mathscr{C}_B(\text{cubes})\}.$$

(*Hint*: Use Problem 3.19.)

3.21 Show that if N is a negligible set in $(\mathbb{R}^n, \mathcal{B}, \lambda)$, for each $\varepsilon > 0$, there is a countable cover of N by disjoint semi-open cubes $\{C_k\}$ such that

$$\sum_{k=1}^{\infty} \lambda^0(C_k) < \varepsilon.$$

3.22 Show that if N is a subset of $\mathbb{R}^n$, and for each $\varepsilon > 0$, there is a countable cover of N by half-open (not necessarily disjoint) cubes such that

$$\sum_{k=1}^{\infty} \lambda^0(C_k) < \varepsilon,$$

then N is negligible.

3.23 Show additivity of f in Theorem 3.10 for the other combinations of x and y.

3.24 Let μ be a translation-invariant Borel measure on $\mathcal{B}(\mathbb{R}^n)$ and let μ^* be the outer measure produced by $(\mathcal{B}(\mathbb{R}^n), \mu)$ and $\mathcal{B}^*_\mu$ be the corresponding σ-algebra of μ^*-measurable sets. Show that

(*i*) $\mu^* = \mu(C)\lambda^*$ on $\mathcal{P}(\mathbb{R}^n)$ and $\mu^*_0 = \mu(C)\lambda^*_0$ on $\mathcal{B}^*_\mu$, where C is a unit cube.

(*ii*) $\mathcal{B}^*_\mu = \mathcal{L}^*$.

3.25 Let $(\mathbb{R}^n, \mathcal{B}, \lambda)$ be the Borel - Lebesgue measure space and let B be a compact set. Show that for each $\varepsilon > 0$, there is a finite cover of B by half-open rectangles $D_1, \ldots, D_N$ such that

$$\sum_{k=1}^{N} \lambda(D_k) - \varepsilon \leq \lambda(B) \leq \sum_{k=1}^{N} \lambda(D_k). \qquad \text{(P3.25)}$$

3.26 For any $\varepsilon > 0$, construct an open set D in $(\mathbb{R}, \tau_e)$, which is dense in $\mathbb{R}$ and with $\lambda(D) \leq \varepsilon$.

3.27 Is every σ-finite Borel measure on $\mathcal{B}(\mathbb{R}^n)$ also a Borel - Lebesgue - Stieltjes measure?

3.28 Show that there is a non-Borel set in $\mathcal{P}(\mathbb{R})$.

NEW TERMS:

λ_c, Lebesgue content, 313
λ^*, Lebesgue outer measure, 316
$\mathcal{L}^*$, Lebesgue σ-algebra, 316
$\lambda_0^* = \text{Res}_{\mathcal{L}^*}\lambda^*$, Lebesgue measure, 316
$(\mathbb{R}^n, \mathcal{L}^*, \lambda_0^*)$, Lebesgue measure space, 316
$\lambda = \text{Res}_{\Sigma(\mathcal{S})}\lambda^* = \text{Res}_{\mathcal{B}(\mathbb{R}^n)}\lambda^*$, Borel - Lebesgue measure, 316
Borel measure, 317
$(\Omega, \mathcal{B}, \mu)$, Borel measure space, 317
Borel - Lebesgue - Stieltjes measure, 317
Borel outer measure, 317
Lebesgue - Stieltjes σ-algebra, 317
Lebesgue - Stieltjes measure, 317
Lebesgue - Stieltjes elementary content, 317
Lebesgue - Stieltjes content, 318
figure, 318
distribution function, 320
extended distribution function, 320
$\mathfrak{D}_e|\text{E}: = \{\{f; \mu\} : \mu \in \mathfrak{B}\}$, quotient class of extended distribution functions, 321
$\mathfrak{B}_*$, finite Borel - Lebesgue - Stieltjes measures, 321
translation-invariant Borel measure, 323
Cantor ternary set, 327-328
μ-outer regular Borel set, 330
μ-inner regular Borel set, 330
μ-closed regular set, 330
outer regular measure, 330
inner regular measure, 330
closed regular measure, 330
Radon measure, 330
regular measure, 330
weakly regular measure, 330
strongly regular measure, 330
measure of a hyperplane, 331
non-Borel set, 334

4. IMAGE MEASURES

In Remark 3.5(iii) we saw how Borel - Lebesgue - Stieltjes measures can be generated by measurable functions belonging to the class of *extended distribution functions* $\mathfrak{D}_e$. We begin this section by generating measures by elements of $\mathcal{C}^{-1}$-class of measurable functions (more general than $\mathfrak{D}_e$) and then illustrate such constructions in Euclidean spaces. The means of generation of measure here is totally different than those in Remark 3.5(iii) and the two notions should not be confused with each other. Section 3, Chapter 4, is a relevant prerequisite to this material.

4.1 Proposition. *Let* (Ω, Σ, μ) *be a measure space and let* $f : (\Omega, \Sigma) \to (\Omega', \Sigma')$ *be a measurable function. Then the set function* $A' \mapsto \mu f^*(A') = \mu(f^*(A'))$ *is a measure on* Σ'. □

(See Problem 4.1.)

4.2 Definition and Notation. The measure μf^* in Proposition 4.1 induced by a measurable function f is called an *image measure.* Notice that directly from Definition 2.1($viii$), Chapter 1, of the inverse image of a set, $\mu f^*(A')$ can alternatively be viewed as $\mu\{\omega \in \Omega\colon f(\omega) \in A'\}$ or, in short, as $\mu\{f \in A'\}$. □

4.3 Proposition. *Let* $L\colon \mathbb{R}^n \to \mathbb{R}^n$ *be such that* $L(x) = \alpha x + b$, *where* $\alpha \in \mathbb{R}\backslash\{0\}$ *and* $b \in \mathbb{R}^n$. *Then the* Borel - Lebesgue *measure* λ *on* $\mathcal{B}(\mathbb{R}^n)$ *has the property* $\lambda L^* = \frac{1}{|\alpha|^n}\lambda$. *Specifically, if* $\alpha = 1$ *we have* $\lambda L^* = \lambda$, *which shows that the* Borel - Lebesgue *measure is translation-invariant*.

Proof.

1) Let $f(x) = \alpha x$ (called a *homothetic function*), where $x \in \mathbb{R}^n$ and $\alpha\,(\neq 0) \in \mathbb{R}$. Let λ be the Borel - Lebesgue measure on the Borel σ-algebra $\mathcal{B}$. We first show that

$$\lambda f^* = \tfrac{1}{|\alpha|^n}\lambda$$

on $\mathcal{S}$. Take $(a, b] \in \mathcal{S}$. Then,

$$f^*((a,b]) = \begin{cases} \prod_{i=1}^{n} \left(\frac{a_i}{\alpha}, \frac{b_i}{\alpha}\right], & \alpha > 0 \\ \prod_{i=1}^{n} \left[\frac{b_i}{\alpha}, \frac{a_i}{\alpha}\right), & \alpha < 0 \end{cases}$$

which implies that

$$\lambda f^*((a,b]) = \tfrac{1}{\alpha^n}\, \lambda((a,b]) \text{ for } \alpha > 0$$

and

$$\lambda f^*((a,b]) = \tfrac{1}{\alpha^n}(-1)^n \lambda((a,b]) \text{ for } \alpha < 0,$$

and thus

$$\lambda f^*((a,b]) = \tfrac{1}{|\alpha|^n} \lambda((a,b]).$$

As a continuous map relative to the usual topology, f is Borel and, thus by Proposition 4.1, λf^* is a Borel measure on $\mathcal{B}$. Obviously, $\frac{1}{|\alpha|^n}\lambda$ is also a Borel measure on $\mathcal{B}$. Because λf^* and $\frac{1}{|\alpha|^n}\lambda$ coincide on $\mathcal{S}$ and clearly are σ-finite and σ-additive on $\mathcal{S}$, by Corollary 2.14, they should also coincide on $\mathcal{B}$.

2) Let $g(x) = x + b$. Similarly, we can show (see Problem 4.2) that λg^* is a Borel measure on $\mathcal{B}$ and furthermore that $\lambda g^* = \lambda$. Therefore, λ is translation-invariant. Finally,

$$L = g \circ f \text{ and } \lambda L^* = \lambda f^* \circ g^* = \tfrac{1}{|\alpha|^n} \lambda g^* = \tfrac{1}{|\alpha|^n} \lambda.$$

The proposition is therefore proved. □

4.4 Remark. From Proposition 4.3 we learn that the Borel - Lebesgue measure is invariant under translation (a particular case of motion) defined as $T_a(x) = x + a$. In the two-dimensional Euclidean space we know another form of the motion, called rotation. A figure under rotation R and subsequent translation T_a is transformed into a congruent one. We will show that an arbitrary Borel set in $\mathbb{R}^n$ "rotated" and then translated preserves its volume [Dsh2].

In the n-dimensional Euclidean space, in place of rotation, we use an orthogonal transformation. More precisely, in $\mathbb{R}^n$ an ortho-

gonal transformation is in the form of an orthogonal $n \times n$ matrix; recall that an $n \times n$ nonsingular matrix A is *orthogonal* if $AA^{\mathrm{T}} = A^{\mathrm{T}}A = I$ (the identity matrix). The composition $M = T_a \circ A$ is an example of a *motion*. Generally, a bijective map M from one metric space (X, d) onto another metric space (Y, ρ) is called an *isometry* if it *preserves the distance*, that is, if for every pair $x, y \in X,\ d(x, y) = \rho(M(x), M(y))$. Two such metric spaces are said to be *isometric*. A *motion* is thus a special case of isometry when $Y = X,\ \rho = d$. Indeed, it is well-known that an orthogonal matrix A preserves the distance ($\|A(x)\|_e = \|Ax\|_e = \|x\|_e$, $x \in \mathbb{R}^n$) and so does the translation. In the Euclidean space $(\mathbb{R}^n, d_e)$, a motion M can be represented by the composition $T_a \circ A,$ where T_a is a translation map and A is an orthogonal matrix. As a continuous map, the motion is also Borel. It can be shown (see Problem 4.10) that the Borel - Lebesgue measure is *motion-invariant*, that is, $\lambda M^* = \lambda$. □

4.5 Examples.

(i) Let (Ω, Σ, μ) be a probability space and let (Ω', Σ') be a measure space. Then any Σ-Σ'-measurable function $f\colon \Omega \to \Omega'$ is called a *random variable* (in short, *r.v.*). The corresponding image measure μf^* is called the *probability distribution* (*of the r.v.* f). Observe that in probability theory, a probability measure is denoted by $\mathbb{P}$ and a random variable is denoted by uppercase letters such as $X, Y,$ or Z. In most applications, Ω' is the numeric set $\mathbb{R}^n$ or a subset of $\mathbb{R}^n$, and Σ' is the corresponding Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ or its trace on the subset. We emphasize that a r.v., say X, is not just a measurable function. It can only be a r.v. if it is associated with a particular probability measure $\mathbb{P}$, along with which it induces the probability distribution. The latter then specifies the r.v.. In other words, measurable functions may share the same measurable space, but as far as probability theory goes, they differ if they generate different probability distributions (or more precisely, different classes of probability distributions categorized by their parameters). Furthermore, as we show later, r.v.'s that induce the same probability distribution may still differ, but they form equivalence classes.

(ii) Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and $X : \Omega \to \{0, 1, \dots\}$ be a r.v. such that $\mathbb{P}X^*$ is a Poisson measure with parameter α:

$$\mathbb{P}X^* = \pi_\alpha = \sum_{n=0}^{\infty} e^{-\alpha} \frac{\alpha^n}{n!} \varepsilon_n \,, \ \alpha > 0.$$

Then the r.v. X is called a *Poisson r.v.*. Similarly, a r.v. X: $\Omega \to \{0, \dots, n\}$ is called *binomial,* if $\mathbb{P}X^*$ is the binomial measure

$$\mathbb{P}X^* = \beta_{n,p} = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \varepsilon_k, \ p \in (0, 1).$$

A r.v. X is called (discrete) *uniformly distributed* if $\mathbb{P}X^* = \sum_{k=1}^{n} \frac{1}{n} \varepsilon_k$. As pointed out in (i), $X : \Omega \to \{0, \dots, n\}$ is just a measurable function (which can be uniform or binomial), and it becomes a r.v. upon specification of its distribution $\mathbb{P}X^*$ or even prior to this, the probability measure $\mathbb{P}$.

These are examples of so-called *discrete r.v.'s*. The construction of probability distributions of *continuous* r.v.'s (i.e., those whose ranges are continuums) requires integration and the concept of a density. The latter is developed in Chapter 6.

(iii) Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and X_i be an (Ω, Σ)-(Ω_i, Σ_i) r.v., $i = 1, \dots, n$. Now, each of the n r.v.'s generates a σ-algebra, $\sigma(X_i) : = X_i^{**}(\Sigma_i) \subseteq \Sigma$. The r.v.'s $X_1, \dots, X_n$ are said to be *independent* if the σ-algebras $\{\sigma(X_1), \dots, \sigma(X_n)\}$ are independent. (For more details on independent families of events, see Problems 1.18 - 1.22 of this chapter. □

PROBLEMS

4.1 Prove Proposition 4.1. *Let (Ω, Σ, μ) be a measure space and let f: $(\Omega, \Sigma) \to (\Omega', \Sigma')$ be a measurable function. Then the set function $A' \mapsto \mu f^*(A') = \mu(f^*(A'))$ is a measure on Σ'.*

4.2 Prove part 2 of Proposition 4.3. *The Borel - Lebesgue measure λ on $\mathcal{B}(\mathbb{R}^n)$ is translation-invariant.*

4.3 Let (Ω, Σ, μ) be a measure space with $\Omega = \mathbb{R}$, $\Sigma = \{A \subseteq \mathbb{R}$: either A or $A^c \preceq \mathbb{N}\}$ and let $\mu(A) = 0$ for $A \preceq \mathbb{N}$ and

$\mu(A) = 1$ for $A^c \preceq \mathbb{N}$. Let $\Omega' = \mathbb{R}$, $\Sigma' = \mathcal{B}(\mathbb{R})$. Define $[\Omega, \Omega', f]$ as

$$f(x) = \mathbf{1}_{\mathbb{Q}^c}(x) = \begin{cases} 0, & \text{if } x \text{ is rational} \\ 1, & \text{if } x \text{ is irrational.} \end{cases}$$

Prove that f is Σ-Σ'-measurable and determine μf^*.

4.4 What are the traces of Borel σ-algebras $\mathcal{B}(\mathbb{R})$ on $\Omega' = \{0, 1, \dots\}$ and $\Omega' = \{0, 1, \dots, n\}$ introduced in Example 4.5 (ii)?

4.5 Let $\mathcal{M}(\mathbb{R}^n)$ be the collection of all motions on $(\mathbb{R}^n, d_e)$. Show that $(\mathcal{M}(\mathbb{R}^n), \circ)$, where $\circ$ is the composition operator, forms a group.

4.6 Let f be a homothetic function $(f(x) = \alpha x)$ defined in Proposition 4.3, part 1, λ be the Borel - Lebesgue measure on $\mathcal{B}(\mathbb{R}^n)$, λ^* be the Lebesgue outer measure, $\mathcal{L}^*$ be the σ-algebra of Lebesgue measurable sets, and λ_0^* be the Lebesgue measure on $\mathcal{L}^*$. Let μ^* be the outer measure generated by the image measure λf^*, $\mathcal{B}_\mu^*$ be the σ-algebra of all μ^*-measurable sets, and $\mu_0^* = \text{Res}_{\mathcal{B}_\mu^*} \mu^*$. Show that:

$a)$ $\mu^* = |\alpha|^{-n} \lambda^*$ on $\mathcal{P}(\mathbb{R}^n)$.

$b)$ $\mu_0^* = |\alpha|^{-n} \lambda_0^*$ on $\mathcal{B}^*$.

$c)$ $\mathcal{B}_\mu^* = \mathcal{L}^*$.

4.7 Generalize Problem 4.6 by letting f be an affine map $f(x) = \alpha x + b$, $\alpha \neq 0$, $b \in \mathbb{R}^n$.

4.8 Show that the Lebesgue measure λ_0^* on $\mathcal{L}^*$ is translation-invariant.

4.9 Let μ be a translation-invariant Borel measure on $\mathcal{B}(\mathbb{R}^n)$ and let μ^* be the outer measure produced by $(\mathcal{B}(\mathbb{R}^n), \mu)$. Show that:

$a)$ $\mu^* = \mu(C)\lambda^*$, where C is the unit cube.

$b)$ $\mathcal{B}^*_\mu = \mathcal{L}^*$.

4.10 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and X_i be a (Ω, Σ)-(Ω_i, Σ_i) r.v., $i = 1, \ldots, n$. Show that $X_1, \ldots, X_n$ are independent if and only if σ-algebras Σ_i's have $\cap$-stable generators $\mathcal{G}_i$ such that the system $\{X_1^{**}(\mathcal{G}_1), \ldots, X_n^{**}(\mathcal{G}_n)\}$ of families of events are independent.

4.11 We know from probability theory that r.v.'s $X_1, \ldots, X_n$ defined on a common probability space $(\Omega, \Sigma, \mathbb{P})$ and valued in $\Omega_1, \ldots, \Omega_n$, endowed with σ-algebras $\Sigma_1, \ldots, \Sigma_n$, are independent, if for any n-tuple of measurable sets $A_1, \ldots, A_n$, $\mathbb{P}\{X_1 \in A_1, \ldots, X_n \in A_n\} = \mathbb{P}\{X_1 \in A_1\} \cdots \mathbb{P}\{X_n \in A_n\}$. How do you reconcile this version of independence with that in Example 4.5 (iii)?

NEW TERMS:

$\mathcal{C}^{-1}$-class of functions, 336
image measure, 336
homothetic function, 336
translation, 337
orthogonal transformation, 338
motion, 338
isometry, 338
isometric metric spaces, 338
motion-invariant measure, 338
random variable, 338
probability distribution, 338
Poisson random variable, 339
binomial random variable, 339
uniform random variable, 339
independent random variables, 339
translation-invariance of Borel - Lebesgue measure, 339
independence criteria of random variables, 341

5. EXTENDED REAL-VALUED MEASURABLE FUNCTIONS

5.1 Definitions and Notations.

(i) Recall (Section 3, Chapter 4) that $\mathcal{C}^{-1}(\Omega, \Sigma; \Omega', \Sigma',)$ denotes the collection of all measurable functions from a measurable space (Ω, Σ) to a measurable space (Ω', Σ'). If $\Omega' = \mathbb{R}$ and $\Sigma' = \mathcal{B}(\mathbb{R})$, then

$$\mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{R}) = \mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{R}, \mathcal{B}(\mathbb{R}))$$

will denote the class of all real-valued measurable functions on a measurable space (Ω, Σ). The class of all complex-valued measurable functions is denoted by

$$\mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{C}) = \mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{C}, \mathcal{B}(\mathbb{C})).$$

Using the notion of product measures (Section 9, Chapter 6) we can show that a function $f = u + iv \in \mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{C})$ if and only if $u, v \in \mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{R})$.

(ii) The "extended real line $\overline{\mathbb{R}}$" that we utilize in this and all forthcoming chapters is formed by a "two-points compactification" procedure whose formalism here is reduced to a minimum.

Firstly, we adhere $\{+\infty\}$ and $\{-\infty\}$ to $\mathbb{R}$ arriving at the extended real line. [See also Example 1.2(iv).] Then, we form the collections of sets $[-\infty, b)$ and $(a, \infty]$ as a subbase for $\overline{\tau}_e$.

The intersections of such sets produce open intervals (a, b), which, along with $[-\infty, a)$ and $(b, \infty]$, according to Theorem 2.8, Chapter 3, form a base for the topology that we denoted by $\overline{\tau}_e$. This topology obviously includes all open sets of τ_e and in addition, open sets of types $[-\infty, b)$ and $(a, \infty]$.

(iii) Now, $(-\infty, b)$ and (a, ∞) are open sets in $\overline{\tau}_e$. It can be readily deduced that the sets $[-\infty, \infty)$ and $(-\infty, \infty]$ are also open and thus the singletons $\{-\infty\}$ and $\{\infty\}$ are closed which comes to

no surprise, since $(\overline{\mathbb{R}}, \overline{\tau}_e)$ is Hausdorff where all singletons are closed sets.

We also notice that closed sets in τ_e need not be closed in $\overline{\tau}_e$. For example, $[a, \infty)$ is a closed set in τ_e but not closed in $\overline{\tau}_e$. This is because $\{\infty\}$ is a closure point of $[a, \infty)$ (because each open neighborhood $(c, \infty]$ of $\{\infty\}$ meets $[a, \infty)$.

(iv) Now, $\overline{\mathbb{R}}$ is compact; firstly, because the spaces $(F, \tau_e \cap F)$, where $F = [-\frac{\pi}{2}, \frac{\pi}{2}]$, and $(\overline{\mathbb{R}}, \overline{\tau}_e)$ are homeomorphic through $[F, \overline{\mathbb{R}}, \text{Res}_F \tan]$. Secondly, we can see that any open cover $\{O_i\}$ of $\overline{\mathbb{R}}$ must include open sets $[-\infty, c)$ and $(d, \infty]$ to cover $\{-\infty\}$ and $\{\infty\}$, respectively. If so, then the set $[c, d]$ (without loss of generality, $c < d$) can be covered by $\{O_i\}$ excluding $[-\infty, c)$ and $(d, \infty]$. Because $[c, d]$ is compact in τ_e, the rest of the cover by $\{O_i\}$ is finitely reducible, thereby making $[-\infty, c)$ and $(d, \infty]$, along with a finite subcover of $\{O_i\}$, a finite subcover $\overline{\mathbb{R}}$.

Like $(\mathbb{R}, \tau_e)$, the compactified space $(\overline{\mathbb{R}}, \overline{\tau}_e)$ is second countable and separable, with $\mathbb{Q}$ being the same dense subset.

We notice that the above formalism of $(\overline{\mathbb{R}}, \overline{\tau}_e)$ can be more rigorously elaborated, but the above is sufficient for the forthcoming applications.

(v) The corresponding Borel σ-algebra $\mathcal{B}(\overline{\mathbb{R}}) = \Sigma(\overline{\tau}_e)$, therefore, consists of all sets of $\mathcal{B}(\mathbb{R})$ and combinations of unions of Borel sets with closed sets $\{+\infty\}$ and $\{-\infty\}$.

In this section, we are concerned with the class of all *extended real-valued* functions $f\colon \Omega \to \overline{\mathbb{R}}$ that are Σ-$\mathcal{B}(\overline{\mathbb{R}})$-measurable, where Σ is a σ-algebra in Ω. We denote such a class by

$$\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}) = \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$$

(or sometimes in short by $\mathcal{C}^{-1}$ if a measurable space (Ω, Σ) is previously specified). □

We begin with a simple criterion for measurability of $\mathcal{C}^{-1}$-functions.

5.2 Proposition. *A function* $f \in \mathcal{C}^{-1}$ *is measurable if and only if, for every real value* a, *the set* $\{\omega \in \Omega\colon f(\omega) \le a\} = f^*([-\infty, a])$, *in notation* $\{f \le a\}$, *is measurable, that is, is an element of* Σ.

Proof. We show that the collection of sets $\{[-\infty, a]\colon a \in \overline{\mathbb{R}}\}$ is a generator of $\mathcal{B}(\overline{\mathbb{R}})$. Then the statement follows directly from Proposition 3.4, Chapter 4.

Denote $\mathcal{B}' = \Sigma(\{[-\infty, a]\colon a \in \overline{\mathbb{R}}\})$. Then,

$$(a, b] = [-\infty, b] \backslash [-\infty, a] \in \mathcal{B}'$$

and in as much as $\mathcal{S} \subseteq \mathcal{B}'$, it immediately implies that $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}'$. Because

$$\{+\infty\} = \bigcap_{k=1}^{\infty} [k, +\infty] \text{ and } \{-\infty\} = \bigcap_{k=1}^{\infty} [-\infty, -k],$$

we have that $\{+\infty\}$, $\{-\infty\} \in \mathcal{B}'$. Thus $\mathcal{B}(\overline{\mathbb{R}}) \subseteq \mathcal{B}'$. On the other hand,

$$[-\infty, a] = \{-\infty\} \cup (-\infty, a) \cup \{a\} \in \mathcal{B}(\overline{\mathbb{R}}).$$

Therefore, $\{[-\infty, a]\colon a \in \overline{\mathbb{R}}\} \subseteq \mathcal{B}(\overline{\mathbb{R}})$, which yields that $\mathcal{B}' \subseteq \mathcal{B}(\overline{\mathbb{R}})$ and, finally, $\mathcal{B}' = \mathcal{B}(\overline{\mathbb{R}})$. □

Proposition 5.2 can be extended to a number of modifications of conditions equivalent to measurability.

5.3 Corollary. *A function* $f \in \mathcal{C}^{-1}$ *is measurable if and only if any of the following conditions holds.*

(i) $\{f \ge a\} \in \Sigma$, $\forall a \in \mathbb{R}$.

(ii) $\{f > a\} \in \Sigma$, $\forall a \in \mathbb{R}$.

(iii) $\{f < a\} \in \Sigma$, $\forall a \in \mathbb{R}$.

Proof. (iii)

$$\{f < a\} = f^*([-\infty, a))$$

$$= f^*\Big(\bigcup_{n=1}^{\infty}[-\infty,\, a-\tfrac{1}{n}]\Big) = \bigcup_{n=1}^{\infty} f^*([-\infty, a-\tfrac{1}{n}])$$
$$= \bigcup_{n=1}^{\infty}\{f \le a-\tfrac{1}{n}\};$$

i.e., $\{f \le b\} \in \Sigma,\ \forall b \in \mathbb{R},$ and thus $\{f < a\} \in \Sigma$. Similarly,

$$\{f \le a\} = f^*\Big[\bigcap_{n=1}^{\infty}[-\infty, a+\tfrac{1}{n})\Big]$$

$$= \bigcap_{n=1}^{\infty} f^*([-\infty, a+\tfrac{1}{n})) = \bigcap_{n=1}^{\infty}\{f < a+\tfrac{1}{n}\},$$

that is, $\{f < b\} \in \Sigma,\ \forall b \in \mathbb{R}$ implies that $\{f \le a\}$.

Therefore, the statements, $\{f \le a\} \in \Sigma,\ \forall a \in \mathbb{R},$ and $\{f < a\}$, $\forall a \in \mathbb{R},$ are equivalent, which along with Proposition 5.2, yield statement (iii). □

[For the proof of (i) and (ii), see Problem 5.1.]

5.4 Proposition. *Let* $f, g \in \mathcal{C}^{-1}$. *Then* $\{f < g\} \in \Sigma$.

Proof. We show that

$$\{f < g\} = \bigcup_{r\in\mathbb{Q}} \{f < r\} \cap \{g > r\}$$

by using the pick-a-point process. We exclude the trivial case when $\{f < g\} = \emptyset$. If $\omega_0 \in \{f(\omega) < g(\omega)\},$ then equivalently, $f(\omega_0) < g(\omega_0),$ implying that there exists an $r_0 \in \mathbb{Q}$ such that $f(\omega_0) < r_0 < g(\omega_0)$. Hence,

$$\omega_0 \in \{f(\omega) < r_0\} \cap \{g(\omega) > r_0\}$$

and

$$\{f < g\} \subseteq \bigcup_{r\in\mathbb{Q}} \{f < r\} \cap \{g > r\}.$$

Conversely, if

$$\omega_0 \in \bigcup_{r\in\mathbb{Q}} \{f < r\} \cap \{g > r\},$$

then there exists an $r_0 \in \mathbb{Q}$ such that $\omega_0 \in \{f(\omega) < r_0\} \cap \{g(\omega > r_0\}$ and $f(\omega_0) < r_0 < g(\omega_0)$, implying that $f(\omega_0) < g(\omega_0)$. Thus $\omega_0 \in \{f(\omega) < g(\omega)\}$. Now the statement follows from Proposition 5.3(ii, iii). □

5.5 Definitions. On some occasions we deal with spaces of measurable functions that have not occurred before. We discuss the following constructions.

(i) Let $\mathbb{F}$ be a field and let $\mathcal{X}$ be a vector lattice over $\mathbb{F}$ and a commutative ring with unity. Observe this makes $(\mathcal{X}, \mathbb{F})$ an algebra and $(\mathcal{X}, \cdot\,)$ a multiplicative Abelian semigroup with unity; call it for short an *$\mathcal{X}$-space over $\mathbb{F}$*. Throughout the remainder of this book, as an $\mathcal{X}$-space, we consider a class of functions (extended real- or complex-valued over the field $\mathbb{R}$ or $\mathbb{C}$). For instance, the space of all continuous functions is an $\mathcal{X}$-space over $\mathbb{R}$. (Note that the term *$\mathcal{X}$-space* is not common in real analysis literature and is restricted to the use in this book.)

(ii) Let $\overline{\mathbb{R}}^{\Omega}$ be the set of all functions from Ω to $\overline{\mathbb{R}}$ (as we defined it in Section 5, Chapter 1), and let $(\overline{\mathbb{R}}^{\Omega}, \tau_P)$ be the topology of pointwise convergence (cf. Definition 5.12, Chapter 3) generated by the compact topology $(\overline{\mathbb{R}}, \overline{\tau}_e)$ in each of the factor spaces. Let us call $(\overline{\mathbb{R}}^{\Omega}, \tau_P)$ the *extended topology of pointwise convergence*. Let Y be a subset of $(\overline{\mathbb{R}}^{\Omega}, \tau_P)$ such that it is an $\mathcal{X}$-space over $\mathbb{R}$. We call Y an *l-$\mathcal{X}$-space* if (Y, τ_P) contains its all sequential limits, that is, the limits of all τ_P-convergent sequences. (Notice that an l-$\mathcal{X}$-space need not be closed.) In other words, it contains the limit of every pointwise convergent sequence of functions. Observe that because by Problem 5.17, Chapter 3, $(\overline{\mathbb{R}}^{\Omega}, \tau_P)$ is Hausdorff and Hausdorffness is hereditary, any pointwise limit in a subspace of $(\overline{\mathbb{R}}^{\Omega}, \tau_P)$ is unique. For instance, the space of all continuous functions is not an l-$\mathcal{X}$-space.

(iii) Consider the subspace $(\mathcal{C}^{-1}, \tau_P) \subseteq (\overline{\mathbb{R}}^{\Omega}, \tau_P)$ of all measurable functions structured in terms of the extended topology of pointwise convergence. The theorem below states that $(\mathcal{C}^{-1}, \tau_P)$ is an l-

$\mathcal{X}$-space and it is the widest known class of functions (second to $\overline{\mathbb{R}}^{\Omega}$). □

5.6 Proposition. *$(\mathcal{C}^{-1}, \tau_P)$ is an l-$\mathcal{X}$-space over $\mathbb{R}$; that is, for any $f, g \in \mathcal{C}^{-1}$ and for $\{f_n : n = 1, 2, \dots\} \subseteq \mathcal{C}^{-1}$:*

(i) $af + b \in \mathcal{C}^{-1}$ $(a, b \in \mathbb{R})$.

(ii) $f \pm g \in \mathcal{C}^{-1}$.

(iii) $f \cdot g \in \mathcal{C}^{-1}$.

(iv) $\sup\{f_n\} \in (\mathcal{C}^{-1}, \tau_P)$ *and* $\inf\{f_n\} \in (\mathcal{C}^{-1}, \tau_P)$; *specifically, it follows that $\mathcal{C}^{-1}$ is a lattice, and thus with any $f \in \mathcal{C}^{-1}$, also $|f| \in \mathcal{C}^{-1}$.*

(v) $\overline{\lim} f_n = \inf_{n=1,2,\dots}\sup\{f_m : m \geq n\} \in (\mathcal{C}^{-1}, \tau_P)$.

$$\underline{\lim} f_n = \sup_{n=1,2,\dots}\inf\{f_m : m \geq n\} \in (\mathcal{C}^{-1}, \tau_P).$$

(vi) *If $f_n \to f$ in the extended topology of pointwise convergence, then $f \in (\mathcal{C}^{-1}, \tau_P)$.*

(vii) *The space $(\mathcal{C}^{-1}, \tau_P)$ is Hausdorff.*

Proof.

(i) Is obvious.

(ii) By (i), $a - g \in \mathcal{C}^{-1}$ implying that $\{f + g \leq a\} = \{f \leq g - a\}$. Therefore, by Problem 5.2 (i), $f + g \in \mathcal{C}^{-1}$.

(iii) Firstly, $\{f^2 \geq a\} = \Omega$, if $a \leq 0$. Then

$$\{f^2 \geq a\} = \{f \geq \sqrt{a}\} \cup \{f \leq -\sqrt{a}\}, \text{ for } a > 0,$$

implying that $\{f^2 \geq a\} \in \Sigma \Rightarrow f^2 \in \mathcal{C}^{-1}$. The statement follows from the representation $f \cdot g = \frac{1}{4}(f + g)^2 - \frac{1}{4}(f - g)^2$.

(iv) We show that

$$\{\sup\{f_n\} \leq a\} = \bigcap_{n=1}^{\infty} \{f_n \leq a\}.$$

$\omega_0 \in \{\sup\{f_n\} \le a\}$ if and only if $\sup\{f_n(\omega_0)\} \le a$ or equivalently $f_1(\omega_0) \le a,\ f_2(\omega_0) \le a, \ldots,\ \Leftrightarrow\ \omega_0 \in \{f_i(\omega) \le a\},\ i = 1, 2, \ldots,$ or, equivalently, $\omega_0 \in \bigcap_{n=1}^{\infty} \{f_n(\omega) \le a\}$. The latter implies that

$$\{\sup\{f_n\} \le a\} = \bigcap_{n=1}^{\infty} \{f_n \le a\}.$$

Let $\{f_n\} \subseteq \mathcal{C}^{-1}$. Then $\{-f_n\} \subseteq \mathcal{C}^{-1}$. The statement follows from

$$\inf\{f_n\} = -\sup\{-f_n\}.$$

Now if $f \in \mathcal{C}^{-1},$ it implies that $|f| = \sup\{f, -f\} \in \mathcal{C}^{-1}$.

(v) This statement follows directly from (iv).

(vi) $\lim_{n \to \infty} f_n = f$ if and only if $\underline{\lim}\ f_n = \overline{\lim}\ f_n = f,$ and the statement follows from (v).

(vii) See Problem 5.9. □

In the sequel, we need the notion of "properties that hold almost everywhere."

5.7 Definitions and Remarks.

(i) Let (Ω, Σ, μ) be a measure space. A property Π (of points of Ω) is said to *hold almost everywhere* $(a.e.)$ or *μ-almost everywhere* (μ-*a.e.*) if there is a (μ-null) set $N \in \mathcal{N}_\mu$ (see Definition 2.5(i), Chapter 5) such that Π holds for all points of N^c. Notice that this definition does not preclude property Π to hold on N or on its subset. It merely says that Π may possibly fail on a negligible subset of N.

(ii) Two functions f and g are said to be *equal* $a.e.$ if $f = g$ on the complement of a μ-null set N. In this case we say that $f = g$ (mod μ). In the event f and g are measurable, by Problem 5.2 (iii), the set $\{f \ne g\}$ is measurable. Therefore, with $f = g$ a.e., the set $\{f \ne g\} \in \mathcal{N}_\mu$ i.e., is μ-null.

In general, however, if $f = g \pmod{\mu}$, where at least one of the functions is not measurable, $\{f \neq g\}$ need not be a null set, although it is negligible, and thus $\{f \neq g\}$ is a subset of a μ-null set.

(iii) Let $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ be the set of all measurable functions on Ω and let μ be a measure on Σ. Let $[f]_\mu$ denote the set of all functions that are pairwise equal μ-a.e. on Ω. Specifically, $[\mathbf{0}]_\mu$ denotes the set of all measurable functions, which equal zero μ-a.e. on Ω. Clearly, the μ-almost everywhere property of equality of functions induces an equivalence relation (say E) on the set $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Then

$$\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})\big|_\mu := \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})\big|_{\mathrm{E}} \tag{5.7}$$

denotes the quotient set $\{[f]_\mu : f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})\}$ and it is called the *quotient set modulo* μ. In light of these considerations, any two functions f and g such that $f = g$ μ-a.e. on Ω are also said to be *equal modulo* μ and we write $f = g \pmod{\mu}$, or $f \in [g]_\mu$, or equivalently, $f - g \in [\mathbf{0}]_\mu$.

(iv) Finally, sometimes we use the notation $[\mathbf{0}]_\mu^+ = [\mathbf{0}]_\mu \cap \mathcal{C}_+^{-1}$, as the subset of all nonnegative measurable functions that equal zero μ-a.e., that we denote

$$\mathcal{C}_+^{-1} = \mathcal{C}^{-1}(\Omega, \Sigma, \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+)). \tag{5.7a}$$

□

PROBLEMS

5.1 Prove Corollary 5.3(i) and (ii).

5.2 Prove that for $f, g \in \mathcal{C}^{-1}$,

(i) $\{f \leq g\} \in \Sigma$.

(ii) $\{f = g\} \in \Sigma$.

(iii) $\{f \neq g\} \in \Sigma$.

5.3 Let $f, g \in \mathcal{C}^{-1}$. Show that $\omega \mapsto \cos(f^2(\omega) + 4g(\omega)) \in \mathcal{C}^{-1}$.

5.4 Show that if $f^3 \in \mathcal{C}^{-1}$ then $f \in \mathcal{C}^{-1}$. Show that if $f^2 \in \mathcal{C}^{-1}$ then f need not be in $\mathcal{C}^{-1}$.

5.5 Let $f, g \in \mathcal{C}^{-1}$ and let $A \in \Sigma$. Show that

$$h = f\mathbf{1}_A + g\mathbf{1}_{A^c} \in \mathcal{C}^{-1}.$$

5.6 Let $f\colon (a, b] \to \overline{\mathbb{R}}$ be $a)$ a monotone function, $b)$ a convex function, or $c)$ a function with at most countably many discontinuities. Show that in each case, f is $\mathcal{B}((a, b])$-$\mathcal{B}(\overline{\mathbb{R}})$-measurable.

5.7 Prove the statement. $f \in \mathcal{C}^{-1}$ if and only if $\{f > q\} \in \Sigma$ for all $q \in \mathbb{Q}$.

5.8 Show that if f has a derivative at each point of $\mathbb{R}$, then this derivative is Borel-measurable.

5.9 Show that the space $(\mathcal{C}^{-1}, \tau_P)$ is Hausdorff.

5.10 Let (Ω, Σ, μ) be a complete measure space and let $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Suppose that $g\colon \Omega \to \overline{\mathbb{R}}$ is an extended, real-valued function. Show that if $g = f \pmod{\mu}$, then $g \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$.

5.11 Let (Ω, Σ, μ) be a complete measure space and let $\{f_n\} \subseteq \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Suppose that $\lim_{n\to\infty} f_n$ exists and $f_n \to f$ pointwise μ-a.e. on Ω, where f is an extended, real-valued function. Show that $f \in \mathcal{C}^{-1}$.

5.12 Let Q be a non-Borel subset of $\overline{\mathbb{R}}$ (such as the one in Problem 3.19) and let C denote the Cantor ternary set. Define the function

$$f(x) = \mathbf{1}_{Q\cap C}(x)\sin x + \mathbf{1}_{(Q\cap C)^c}(x)x^2.$$

Is f Lebesgue measurable, that is, $f \in \mathcal{C}^{-1}(\mathbb{R}, \mathcal{L}^*, \lambda)$?

5.13 We knew from probability theory that r.v.'s $X_1, \ldots, X_n$ on a probability space $(\Omega, \Sigma, \mathbb{P})$ and valued in $\overline{\mathbb{R}}$ (each), endowed with the Borel σ-algebra $\mathcal{B}(\overline{\mathbb{R}})$, are independent if and only if their joint probability distribution function (commonly abbreviated as the PDF) $\mathbb{P}\{X_1 \leq x_1, \ldots, X_n \leq x_n\}$ equals the product $\prod_{i=1}^{n} \mathbb{P}\{X_i \leq x_i\}$ of their marginal PDF's. Prove that this is true using the original definition in Example 4.5(iii).

5.14 Prove the statement: $f \in \mathcal{C}^{-1}$ if and only if $\{f > d\} \in \Sigma$ for all $d \in D$, where $D \subseteq \mathbb{R}$ is any dense set in $\mathbb{R}$.

NEW TERMS:

$\mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{R}) = \mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{R}, \mathcal{B}(\mathbb{R}))$, class of all real-valued measurable functions, 343
$\mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{C}) = \mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{C}, \mathcal{B}(\mathbb{C}))$, class of all complex-valued measurable functions, 343
extended real line, 343
two-points compactification, 343
$\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}) = \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, class of extended real-valued measurable functions, 344
measurability of an extended real-valued function, criteria of, 345
$\mathcal{X}$-space, 347
$(\overline{\mathbb{R}}^{\Omega}, \tau_P)$, extended topology of pointwise convergence, 347
l-$\mathcal{X}$-space, 347
μ-a.e. property, 349
a.e. property, 349
$\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})|_{\mu}$, quotient set modulo μ, 350
$[\mathbf{0}]_{\mu}$, set of all measurable functions equal zero μ-a.e., 350
$[\mathbf{0}]_{\mu}^{+} = [\mathbf{0}]_{\mu} \cap \mathcal{C}_{+}^{-1}$, 350
$\mathcal{C}_{+}^{-1} = \mathcal{C}^{-1}(\Omega, \Sigma, \overline{\mathbb{R}}_{+}, \mathcal{B}(\overline{\mathbb{R}}_{+}))$, 350
independence criterion of numeric random variables, 352

6. SIMPLE FUNCTIONS

The present section is a direct precursor to integration, which we develop in the next chapter. The integral itself is first defined for simple functions valued in a finite set of nonnegative reals.

6.1 Definition. We consider the following subclass of functions from $\mathcal{C}^{-1}(\Omega, \Sigma; \bar{\mathbb{R}})$, which we call *nonnegative simple functions* and denote this subclass by $\Psi_+(\Omega, \Sigma) = \Psi_+(\Omega, \Sigma; \mathbb{R})$. An element s is said to belong to Ψ_+ or to be *nonnegative simple* if:

$a)$ $s \in \mathcal{C}^{-1}$.

$b)$ $s(\omega) \geq 0, \ \forall \omega \in \Omega$.

$c)$ s takes on only finitely many real values. □

6.2 Remarks.

(i) Let $s \in \Psi_+(\Omega, \Sigma)$. If there is an n-tuple of nonnegative real numbers $\{a_1, \ldots, a_n\}$ and a finite decomposition $\sum_{k=1}^n A_k$ of Ω such that $s(\omega) = a_k$ for all $\omega \in A_k$, then the function s can obviously be represented as

$$s = \sum_{k=1}^n a_k \mathbf{1}_{A_k}. \tag{6.2}$$

(See Figure 6.1.)

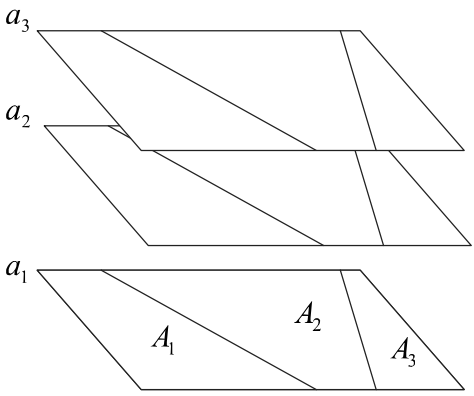


Figure 6.1

In the forthcoming constructions we deal with different partitions of Ω. Consequently, there are different finite representations or expansions of $s \in \Psi_+$ of type (6.1). However, there is obviously a unique one where (6.2) contains all different values $\{a_1, \dots, a_n\}$ of s. We call such a representation (expansion) *canonic.*

(ii) We call the pair $(\mathcal{X}, \mathbb{F}_+)$ a *semi-$\mathcal{X}$-space* if it meets all specifications in Definition 5.5(i), except that $\mathcal{X}$ is defined over a semifield $\mathbb{F}_+$. Recall that $\mathbb{F}_+ \subseteq \mathbb{F}$ is called the *semifield* if all axioms of the field hold except for No. 4 (the existence of additive inverses; see Definition 7.5, Chapter 1).

(iii) In [Dsh2], we also use the notion of a *simple function*, which is just as in Definition 6.1, except that they are not necessarily nonnegative. The set of all such simple functions is denoted by $\Psi(\Omega, \Sigma) = \Psi(\Omega, \Sigma; \mathbb{R})$. □

6.3 Proposition. $(\Psi_+(\Omega, \Sigma)\,;\, \mathbb{R}_+;\, \cdot\,)$ *is a semi-$\mathcal{X}$-space. In other words, if $s, t \in \Psi_+$, then*:

(i) $as + bt \in \Psi_+$ $(a, b \in \mathbb{R}_+)$.

(ii) $s \cdot t \in \Psi_+$.

(iii) $\sup\{s, t\} \in \Psi_+$.

(iv) $\inf\{s, t\} \in \Psi_+$. □

(See Problem 6.1.)

We denote by $(\widetilde{\Psi}_+(\Omega, \Sigma), \tau_P)$ the subspace of all extended, real-valued, nonnegative functions $f \in \mathcal{C}^{-1}$ for each of which there exists a monotone nondecreasing sequence $\{s_n\} \subseteq \Psi_+$ of nonnegative simple functions such that $f = \sup\{s_n\}$ in τ_P. By Proposition 5.5 (iv), $\widetilde{\Psi}_+ \subseteq \mathcal{C}^{-1}$, that is, $\widetilde{\Psi}_+$ consists of only measurable functions. The theorem below asserts that the set of all nonnegative measurable functions $\mathcal{C}_+^{-1}$ coincides with $\widetilde{\Psi}_+$.

6.4 Theorem. *Let $\mathcal{C}_+^{-1} := \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}_+)$, i.e., the subclass of all nonnegative extended real-valued functions. Then for every $f \in \mathcal{C}_+^{-1}$, there is a monotone nondecreasing sequence $\{s_n\}$ of nonnegative simple functions from Ψ_+ such that* $\sup\{s_n\} = f$, *that is,* $\mathcal{C}_+^{-1} = \widetilde{\Psi}_+$.

Proof. As already mentioned, $\widetilde{\Psi}_+ \subseteq \mathcal{C}_+^{-1}$. Therefore, we are left to prove that $\mathcal{C}_+^{-1} \subseteq \widetilde{\Psi}_+$. We show that, for every $f \in \mathcal{C}_+^{-1}$, there is a monotone nondecreasing sequence $\{s_n\}$ of nonnegative simple functions from Ψ_+ such that $\sup\{s_n\} = f$. The latter is at the heart of the following construction. Let

$$A_i(n) =$$

$$\begin{cases} \left\{f > \frac{i}{2^n}\right\} \cap \left\{f \le \frac{i+1}{2^n}\right\}, & i = 0, 1, \dots, n2^n - 1 \\ \{f > n\}, & i = n2^n. \end{cases}$$

For instance,

$$A_0(n) = \{f > 0\} \cap \left\{f \le \tfrac{1}{2^n}\right\} = f^*((0, 2^{-n}]),$$

$$A_1(n) = \left\{f > \tfrac{1}{2^n}\right\} \cap \left\{f \le \tfrac{2}{2^n}\right\} = f^*((2, {}^{-n}, 2^{n-1}]).$$

In other words,

$$A_i(n) = f^*\big(\big(\tfrac{i}{2^n}, \tfrac{i+1}{2^n}\big]\big),\ i = 0, \dots, n2^n - 1,$$

and

$$A_i(n) = f^*((n, \infty]),\ i = n2^n.$$

Therefore, all sets $A_i(n),\ i = 0, \dots, n2^n,$ are disjoint and obviously Σ-measurable. Let us define

$$s_n(\omega) = \sum_{i=0}^{n2^n} i2^{-n} \mathbf{1}_{A_i(n)}(\omega)\ (\in \Psi_+).$$

Both f and s_n are depicted in Figure 6.2.

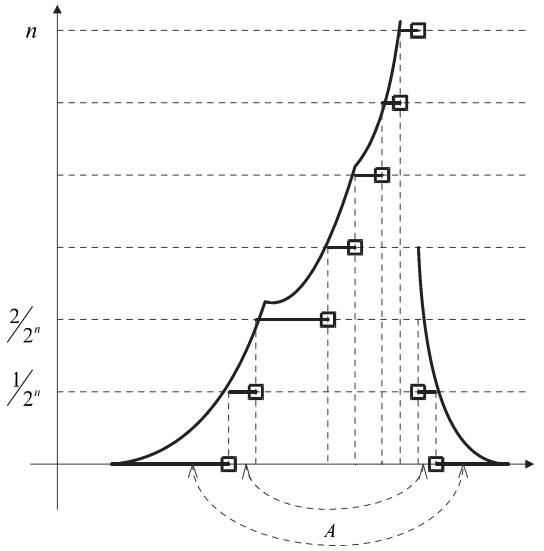


Figure 6.2

Clearly $s_{n+1} \geq s_n$. In addition, $s_n(\omega) \leq f(\omega) < s_n(\omega) + 2^{-n}$, $\forall \omega \in \Omega$: $f(\omega) \leq n$, and $f(\omega) > n$, $\forall \omega \in \Omega$: $f(\omega) = \infty$. Functions s_n and s_{n+1} are drawn in Figure 6.3.

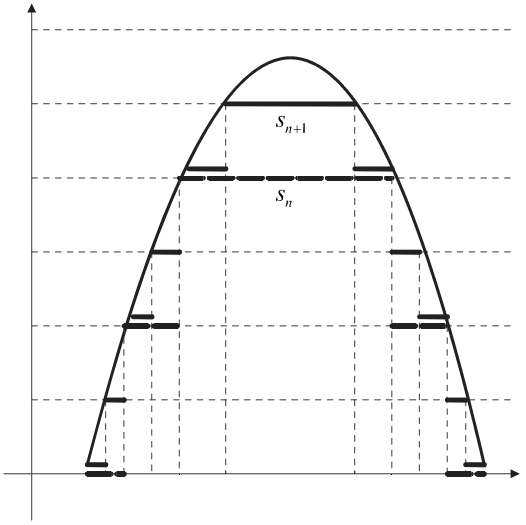


Figure 6.3

Thus there exists $\sup\{s_n\} = f$ (pointwise $\forall\omega \in \Omega$), and therefore $f \in \widetilde{\Psi}_+$, implying that $\mathcal{C}_+^{-1} \subseteq \widetilde{\Psi}_+$.

This proves that $\mathcal{C}_+^{-1} = \widetilde{\Psi}_+$. □

Remark 6.5. Theorem 6.4 can be interpreted as follows. For each $f \in \mathcal{C}_+^{-1}$, there is a sequence $\{s_n\} \subseteq \Psi_+$ such that $s_n \to f$ in τ_P. (That the sequence is monotone increasing is essential only for the existence of such a sequence through an explicit construction.) By Theorem 5.13, Chapter 3, it means that for each base neighborhood B_f, there is an $N(B_f)$-tail of $\{s_n\}$. Consequently, each B_f meets Ψ_+ and hence, $f \in \overline{\Psi}_+$, that is, f belongs to the closure of Ψ_+ in the sense of the Tychonov topology τ_P. We shows that $\mathcal{C}_+^{-1} \subseteq \overline{\Psi}_+$.

To prove $\overline{\Psi}_+ \subseteq \mathcal{C}_+^{-1}$ requires that $\mathcal{C}_+^{-1}$ be first countable, which $\mathcal{C}_+^{-1}$ would inherit from first countability of $\overline{\mathbb{R}}^{\Omega}$ should the latter be first countable to begin with. Unfortunately $\overline{\mathbb{R}}^{\Omega}$ is not first countable, unless Ω is at most countable (a not very practical special case), as we learn from Problems 5.18 and 5.20, Chapter 3. First countability would provide a countable monotone decreasing sequence of base neighborhoods, from which we could select a sequence $\{s_n\}$ convergent to a function φ arbitrarily picked out from the closure $\overline{\Psi}_+$. This would prove that φ is measurable and thus $\overline{\Psi}_+ \subseteq \mathcal{C}_+^{-1}$. Then $\overline{\Psi}_+ = \mathcal{C}_+^{-1}$ would be a desirable property that $\mathcal{C}_+^{-1}$ and thus $\mathcal{C}^{-1}$ were topologically closed.

These all remain wishful thinking though. Recall that in Remark 1.15(ii), Chapter 3, we brought up the case of the cocountable topological space (being Hausdorff and not first countable), in which a set contained all of its sequential limit points and yet failed to be closed. As the result of the above considerations, $\mathcal{C}_+^{-1}$ remains a proper subset of $\overline{\Psi}_+$ (unless Ω is countable). Obviously, $\mathcal{C}_+^{-1}$ is not closed in τ_P, because if it were, we would have $\overline{\Psi}_+ \subseteq \overline{\mathcal{C}}_+^{-1} = \mathcal{C}_+^{-1}$ (which is absurd) for $\Psi_+ \subseteq \mathcal{C}_+^{-1}$. □

6.6 Notation. Now we further enlarge the class of measurable functions by considering arbitrary extended, real-valued, measurable functions of $\mathcal{C}^{-1}(\Omega, \Sigma)$. For each $f \in \mathcal{C}^{-1}$ and $\mathbf{0}$, being the function identically equal to zero on Ω, denote

$$f^+ = \sup\{f, \mathbf{0}\} \quad \text{and} \quad f^- = -\inf\{f, \mathbf{0}) = (-f)^+$$

(cf. Definition 7.7, Chapter 1). Clearly (see also Problem 7.15, Chapter 1),

$$f^+ \geq 0,\ f^- \geq 0,\ f = f^+ - f^-,\ \text{and}\ |f| = f^+ + f^-.$$

By Proposition 6.4, f^+ and f^- are also elements of $\mathcal{C}^{-1}$ (more precisely, elements of $\mathcal{C}_+^{-1}$) if and only if $f \in \mathcal{C}^{-1}$. We revisit this in Chapter 6, Section 1, regarding integration. □

PROBLEMS

6.1 Prove Proposition 6.3.

6.2 Let Ω be an uncountable set and let $\Sigma = \{A \subseteq \Omega\colon\ A \text{ or } A^c \text{ is at most countable}\}$. Show that $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$ if and only if f is constant everywhere except on an at most countable subset of Ω.

6.3 Prove that $f = g \pmod{\mu}$ if and only if $f^+ = g^+ \pmod{\mu}$ and $f^- = g^- \pmod{\mu}$.

6.4 Show that $f \in [\mathbf{0}]_\mu$ if and only if $f^+, f^- \in [\mathbf{0}]_\mu^+$, where $[\mathbf{0}]_\mu^+ = [\mathbf{0}]_\mu \cap \mathcal{C}_+^{-1}$.

6.5 Theorem 6.4 states that for each $f \in \mathcal{C}_+^{-1}$, there is a monotone nondecreasing sequence $\{s_n\}$ of nonnegative simple functions (from Ψ_+) such that $\sup\{s_n\} = f$. In Remark 6.5, it is shown that unless $\mathcal{C}_+^{-1}$ is first countable in the sense of τ_P (which in most applications it is not), $\mathcal{C}_+^{-1}$ is a proper subset of $\bar{\Psi}_+$ (the τ_P-closure of Ψ_+). Does it mean that the converse of Theorem 6.4 does not hold [i.e., that, although $\sup\{s_n\} \in \bar{\Psi}_+$ of some sequence $\{s_n\}$ of simple functions, it need not belong to $\mathcal{C}_+^{-1}$]? Analyze and explain this.

NEW TERMS:

$\Psi_+(\Omega, \Sigma) = \Psi_+(\Omega, \Sigma; \mathbb{R})$, nonnegative simple functions, 354
canonic representation (expansion), 355
canonic expansion (representation), 355
$(\mathcal{X}, \mathbb{F}_+)$, semi-$\mathcal{X}$-space, 355
$\Psi(\Omega, \Sigma) = \Psi(\Omega, \Sigma; \mathbb{R})$, simple functions, 355
$(\widetilde{\Psi}_+(\Omega, \Sigma), \tau_P)$, extended, real-valued, nonnegative functions and their limits, 355
$f^+ = \sup\{f, \mathbf{0}\}$, 359
$f^- = -\inf\{f, \mathbf{0}\}$, 359

Chapter 6

Integration in Abstract Spaces

The historical significance of the development of measure theory is that it created a base for a generalization of the classical Riemann notion of the definite integral (which since 1854 has been considered to be the most general theory of integration). Riemann defined a bounded function over an interval $[a, b]$ to be integrable if and only if the Darboux (or Cauchy) sums $\sum_{i=1}^{n} f(t_i)\lambda(I_i)$, where $\sum_{i=1}^{n} I_i$, is a finite decomposition of $[a, b]$ into subintervals, approach a unique limiting value whenever the length of the largest subinterval goes to zero. The French mathematician, Henri Lebesgue (1875 – 1941), assumed that the above subintervals I_i may be substituted by more general measurable sets and, in addition, that the class of Riemann integrable functions can be enlarged to the class of measurable functions. In this case, we arrive at a more advanced theory of integration, which is better suited for dealing with various limit processes and which led to the contemporary theory of probability and stochastic processes.

Although many results existed prior to Lebesgue's major work between 1901 and 1910, Lebesgue's construction appeared to be the most efficient. After 1910, a large number of mathematicians began to engage in work initiated by Lebesgue. Some of the most significant contributions were made by the Frenchman Pierre Fatou (1878 – 1929), Italian Guido Fubini (1897 – 1943), Hungarian Frigyes Riesz (1880 – 1956), Pole Otto Nikodym (1887 – 1974), and the Austrian Johann Radon (1887 – 1956) who introduced the Lebesgue - Stieltjes integral and whose work led to the modern abstract theory of measure and integration.

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0_6, © Springer Science+Business Media, LLC 2013

We begin this chapter with conventional integration relative to positive measures using the same techniques as those for the original Lebesgue integral (but with no sacrifice to generality). It is followed by applications of the integral (such as those to probability and to the relationship with the Riemann integral in Euclidean spaces). We note that the classes of integrable functions we are concerned with have no assumptions on the topological structure of sets (that is, carriers) on which they are defined. Furthermore, we introduce integrals of these functions also with respect to signed and complex measures. These features warrant the title of this chapter. (In the literature on measure theory, *abstract integration* is sometimes used in the context of more specialized or structured spaces.) We then further proceed with absolute continuity of measures (including signed and complex measures) and multiple integration.

1. INTEGRATION ON $\mathcal{C}^{-1}(\Omega, \Sigma)$

We begin the theory of integration with integrals of nonnegative simple functions, which we introduced in section 6, Chapter 5. Prior to the definition of the rudimentary integral, the proposition below states that integrals of nonnegative simple functions are invariant of their representations.

1.1 Lemma. *Let* (Ω, Σ, μ) *be a measure space and let* $s \in \Psi_+(\Omega, \Sigma)$ *have two representations*:

$$s = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} = \sum_{k=1}^{m} b_k \mathbf{1}_{B_k}.$$

Then it holds that

$$\sum_{i=1}^{n} a_i \mu(A_i) = \sum_{k=1}^{m} b_k \mu(B_k).$$

Proof. The above representations are due to the two decompositions of Ω:

$$\Omega = \sum_{i=1}^{n} A_i = \sum_{k=1}^{m} B_k \,.$$

Then

$$\mu(A_i) = \sum_{k=1}^{m} \mu(A_i \cap B_k) \text{ and } \mu(B_k) = \sum_{i=1}^{n} \mu(A_i \cap B_k),$$

which implies that

$$\sum_{i=1}^{n} \mu(A_i) = \sum_{i=1}^{n} \sum_{k=1}^{m} a_i \mu(A_i \cap B_k)$$

and

$$\sum_{k=1}^{m} b_k \mu(B_k) = \sum_{k=1}^{m} \sum_{i=1}^{n} b_k \mu(A_i \cap B_k).$$

By noticing that $a_i = b_k$ on $A_i \cap B_k$, we are done with the proof. □

1.2 Definition. Let (Ω, Σ, μ) be a measure space and let $s \in \Psi_+(\Omega, \Sigma)$ with the representation

$$s = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}.$$

Then the number

$$\sum_{i=1}^{n} a_i \mu(A_i)$$

is called the *integral of* s *with respect to* μ, and it is denoted by one of the symbols:

$$\int s(\omega) d\mu(\omega) \text{ or } \int s(\omega) \mu(d\omega) \text{ or, for short, } \int s d\mu. \qquad \square$$

The value of the integral of a function s does not depend upon its representation. Consequently, by Lemma 1.1, this definition is consistent. In other words, the integral $s \mapsto \int s d\mu$ defines a functional on Ψ_+ valued in $\overline{\mathbb{R}}$.

1.3 Proposition (Properties of the integral).

(i) *For each measurable set* $A \in \Sigma$,

$$\int \mathbf{1}_A \, d\mu = \mu(A).$$

(ii) *The integral* $\int$ *is a linear functional valued in* $\overline{\mathbb{R}}_+$, *i.e.*,

$$\int (as + bt)\, d\mu = a \int s d\mu + b \int t d\mu,$$

where $s, t \in \Psi_+$ *and* $a, b \in \mathbb{R}_+$.

(iii) *For any two nonnegative simple functions,* $s,\ t \in \Psi_+$, *such that* $s \le t$, *it holds that*

$$\int s d\mu \le \int t d\mu \ (\textit{monotonicity}). \qquad \square$$

(See Problem 1.1.)

1.4 Example. Let f be the *Dirichlet function* defined as $f = \mathbf{1}_{\mathbb{Q}}$ (earlier introduced in Example 4.7, Chapter 2), where $\mathbb{Q}$ is the set of all rational numbers (hence a Borel set). Thus $f \in \Psi_+(\mathbb{R}, \mathcal{B})$. By Proposition 1.3 (i), the integral of f with respect to the Lebesgue measure is

$$\int \mathbf{1}_{\mathbb{Q}} d\lambda = \lambda(\mathbb{Q}) = \sum_{x \in \mathbb{Q}} \lambda(x) = 0. \qquad \square$$

For the upcoming definitions and statements we denote a *monotone nondecreasing sequence of functions* by $\{f_n\}\uparrow$ and a *monotone nonincreasing sequence of functions* by $\{f_n\}\downarrow$.

1.5 Lemma. *Let* $\{s_n\}\uparrow \subseteq \Psi_+$ *and* $s \in \Psi_+$ *such that* $s \le \sup\{s_n\}$. *Then*

$$\int s d\mu \le \sup\{\int s_n d\mu\}.$$

Proof. Let $s = \sum_{i=1}^m a_i \mathbf{1}_{A_i}$ and let $\varepsilon > 0$ be any small number. Denote

$$B_n = \{\omega\colon\ s_n \ge (1 - \varepsilon)s\}\ (\in \Sigma).$$

Thus $s_n \ge s(1 - \varepsilon)\mathbf{1}_{B_n}$. By Proposition 1.3$(ii, iii)$,

$$\int s_n d\mu \ge (1 - \varepsilon) \int s \mathbf{1}_{B_n} d\mu,\ n = 1, 2, \dots\ .$$

By the definition of $\{B_n\}$, it follows that $\{B_n\}\uparrow \Omega$, which implies that $\{A_j \cap B_n\}\uparrow A_j$. Therefore, by continuity from below of μ (Lemma 1.6, Chapter 5),

$$\int s\, d\mu = \sum_{i=1}^{m} a_i \mu(A_i) = \sum_{i=1}^{m} a_i \lim_{n \to \infty} \mu(A_i \cap B_n)$$

$$= \lim_{n \to \infty} \sum_{i=1}^{m} a_i \mu(A_i \cap B_n) = \lim_{n \to \infty} \int s \mathbf{1}_{B_n} d\mu.$$

The last equation is due to the relationship

$$s\mathbf{1}_{B_n} = \mathbf{1}_{B_n} \sum_{i=1}^{m} A_i \mathbf{1}_{A_i} = \sum_{i=1}^{m} a_i \mathbf{1}_{A_i \cap B_n}.$$

Thus,

$$\sup\{\int s_n d\mu)\} = \lim_{n \to \infty} \int s_n d\mu$$

$$\geq (1 - \varepsilon) \lim_{n \to \infty} \int s\, \mathbf{1}_{B_n} d\mu = (1 - \varepsilon) \int s\, d\mu,$$

which proves the statement because the inequality holds for each $\varepsilon > 0$. □

1.6 Corollary. *For* $\{s_n\} \uparrow$, $\{t_n\} \uparrow \subseteq \Psi_+$ *such that* $\sup\{s_n\} = \sup\{t_n\}$, *it holds that*

$$\sup\{\int s_n d\mu\} = \sup\{\int t_n d\mu\}.$$ □

(See Problem 1.2.)

Let us now turn to the integral of the functions from the more general class $\mathcal{C}_+^{-1} = \mathcal{C}^{-1}(\Omega, \Sigma; \bar{\mathbb{R}}_+)$ which we became more familiar with in Theorem 6.4, Chapter 5.

1.7 Definition. Let (Ω, Σ, μ) be a measure space and let $f \in \mathcal{C}_+^{-1}$. By Theorem 6.4, Chapter 5, there is a monotone nondecreasing sequence $\{s_n\} \uparrow \subseteq \Psi_+$ such that $f = \sup\{s_n\}$. Hence, it is plausible to define

$$\int f d\mu = \sup\{\int s_n d\mu\}$$

and call it the *integral of* (*an extended*, *real-valued*, *nonnegative function*) f *with respect to measure* μ. By Corollary 1.6, the value of the integral $\int f d\mu$ is unique. □

Analogous to Proposition 1.3 (ii, iii), we have the following.

1.8 Proposition. *The integral introduced in Definition* 1.7 *is a positive*, *linear*, *monotone nondecreasing functional on* $\mathcal{C}_+^{-1}$.

Proof. Let $f, g \in \mathcal{C}_+^{-1}$ and $a, b \in \mathbb{R}_+$. Then

$$f = \sup\{s_n\},\ g = \sup\{t_n\} \text{ and } af + bg = \sup\{as_n + bt_n\}$$

yield that

$$\int (af + bg)d\mu = \sup\{\int (as_n + bt_n)d\mu\},$$

which, by Proposition 1.3 (ii), equals

$$\sup\{a\int s_n d\mu + b\int t_n d\mu\}$$

$$= a\sup\{\int s_n d\mu\} + b\sup\{\int t_n d\mu\} = a\int f d\mu + b\int g d\mu.$$

Now let $f \leq g$. Then we have:

$$\sup s_n \leq \sup t_n;\ \text{ hence } s_k \leq \sup t_n\,.$$

Thus, by Lemma 1.5,

$$\int s_k d\mu \leq \sup\int t_n d\mu,$$

and finally,

$$\int f d\mu = \sup\int s_k d\mu \leq \int t_n d\mu = \int g d\mu. \qquad \square$$

1.9 Examples.

(i) Let ε_b be a point mass on a measurable space (Ω, Σ) for a fixed $b \in \Omega$ and let $s = \sum_{i=1}^m \alpha_i \mathbf{1}_{A_i} \in \Psi_+(\Omega, \Sigma)$ be such that $s(b) = \alpha_{i_0}$, for some $i_0 \in \{1, \ldots, m\}$. Then

$$\int s d\varepsilon_b = \sum_{i=1}^m \alpha_i \varepsilon_b(A_i) = \alpha_{i_0} \cdot 1 = s(b).$$

Now let $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$. Then there is a sequence $\{s_n\}\uparrow \subseteq \Psi_+$ such that $f = \sup\{s_n\}$. Thus

$$\int f d\varepsilon_b = \sup\{s_n(b)\} = f(b).$$

Moreover, if $\mu = c\varepsilon_b$ (for some $c > 0$), using similar arguments we get $\int f d\mu = cf(b)$.

(ii) Let

$$\mu = \sum_{i=0}^{n} c_i \varepsilon_{a_i}.$$

By Problem 1.4,

$$\int f d\mu = \sum_{i=0}^{n} c_i f(a_i).$$

Specifically, if $c_i = \binom{n}{i} p^i (1-p)^{n-i}$, then μ is the binomial measure $\beta_{n,p}$. [See Example 1.8(iii), Chapter 5.] Furthermore, if $f(x) = e^{\theta x}$, for $\theta \in \mathbb{R}$, then the transform of the binomial measure

$$\theta \mapsto \int f(x)\mu(dx) = \int e^{\theta x} \beta_{n,p}(dx)$$

$$= \sum_{i=0}^{n} c_i e^{\theta i} = (1 + pe^{\theta} - p)^n$$

is a function in θ and is referred to as the *moment generating function*. In the general definition of a moment generating function, θ runs the complex plane $\mathbb{C}$.

(iii) Let (Ω, Σ, μ) be the measure space with $\Omega = [0,1]$, $\Sigma = \mathcal{B}([0,1])$, and $\mu = \lambda$ (Borel - Lebesgue measure on $\mathcal{B}([0,1])$. Let C be the Cantor set and G_n be the open intervals of the Borel - Lebesgue measure $\frac{1}{2}(\frac{2}{3})^n$ (introduced in Example 3.11, Chapter 5). Let us define the function

$$f(x) = \begin{cases} 1, & x \in C \\ \frac{1}{2^n}, & x \in G_n,\ n = 1, 2, \dots \end{cases}.$$

We evaluate the integral $\int\limits_{[0,1]} f(x)\lambda(dx)$ (with respect to the Borel - Lebesgue measure). First of all, we have to identify the function f, which can be represented in the form $f = \sup\{s_n\},$ where

$$s_n(x) = \begin{cases} 1, & x \in C \\ 0, & x \in [0,1]\backslash(G_1 \cup \ldots \cup G_n \cup C) \\ \frac{1}{2^k}, & x \in G_k\,,\ k = 1, \ldots, n\,. \end{cases}$$

Clearly, $s_n \in \Psi_+([0,1],\ \mathcal{B} \cap [0,1])$ and $f(x) = \sup\{s_n(x)\}$. Thus

$$f \in \mathcal{C}_+^{-1}([0,1], \mathcal{B}([0,1]))$$

and hence

$$\int\limits_{[0,1]} f(x)\lambda(dx) = \sup \int\limits_{[0,1]} s_n(x)\lambda(dx)$$

$$= \sup\left[1 \cdot \lambda(C) + 0 \cdot \lambda([0,1]\backslash\{G_1, \ldots, G_n, C\}) + \sum_{k=1}^{n} \tfrac{1}{2^k}\,\lambda(G_k)\right]$$

$$= \sum_{k=1}^{\infty} \tfrac{1}{2}(\tfrac{1}{3})^k = \tfrac{1}{4}.$$

(iv) Let $\{\mu_n\}$ be a sequence of measures on a measurable space (Ω, Σ). Then $\mu = \sum_{n=1}^{\infty}\mu_n$ is a measure on Σ; and for an $A \in \Sigma$, the integral of the indicator function $\mathbf{1}_A$ is

$$\int \mathbf{1}_A\, d\mu = \mu(A) = \textstyle\sum_{n=1}^{\infty}\mu_n(A) = \sum_{n=1}^{\infty}\int \mathbf{1}_A\, d\mu_n.$$

Let $s \in \Psi_+(\Omega, \Sigma)$. Then

$$\begin{aligned} \int s d\mu &= \textstyle\sum_{k=1}^{m} a_k\, \mu(A_k) \\ &= \textstyle\sum_{k=1}^{m} a_k \sum_{n=1}^{\infty}\mu_n(A_k) \qquad (1.9) \\ &= \textstyle\sum_{n=1}^{\infty}\sum_{k=1}^{m} a_k\mu_n(A_k) = \sum_{n=1}^{\infty}\int s d\mu_n. \end{aligned}$$

Now, for $f \in \mathcal{C}_+^{-1}$, we have $f = \sup\{s_j\}$ such that $\{s_j\}\uparrow\ \subseteq \Psi_+$. Let

$$b_{j_n} = \sum_{i=1}^{n} \int s_j d\mu_i.$$

In as much as $\{b_{j_n}\}$ is monotone increasing,

$$\sup_j \sup_n \{b_{j_n}\} = \sup_n \sup_j \{b_{j_n}\},$$

which yields that

$$\sup_j \sum_{i=1}^{\infty} \int s_j \, d\mu_i = \sup_n \sum_{i=1}^{n} \int f d\mu_i = \sum_{i=1}^{\infty} \int f d\mu_i \,. \quad (1.9a)$$

Therefore,

$$\int f d\mu = \sup_j \int s_j d\mu \text{ [by (1.9)]} = \sup_j \sum_{i=1}^{\infty} \int s_j d\mu_i \text{ [(by (1.9a)]}$$

$$= \sum_{i=1}^{\infty} \int f d\mu_i.$$

Thus we showed that

$$\int f d \left(\sum_{n=1}^{\infty} \mu_n \right) = \sum_{n=1}^{\infty} \int f d\mu_n \,, \quad \forall f \in \mathcal{C}_+^{-1}(\Omega, \Sigma). \qquad \square$$

Now we further enlarge the class of integrable functions in light of Notation 6.6, Chapter 5.

1.10 Definitions.

(i) Let (Ω, Σ, μ) be a measure space and let $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ [or $\mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{R})$]. If at least one of the integrals, $\int f^+ d\mu$ or $\int f^- d\mu$, is finite, we say that the *integral of* f *with respect to measure* μ *exists* and denote this integral by

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu. \quad (1.10)$$

We also denote

$$\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}) = \{f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}) : \int f d\mu \text{ exists}\}. \quad (1.10a)$$

If either integral of the functions f^+ and of f^- is finite, we say that the function f is μ-*integrable* and again denote the integral of f by formula (1.10). The subset of $\mathcal{C}^{-1}$ of all μ-integrable functions is denoted by $L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, that is,

$$L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$$

$$= \{f \in \mathcal{C}^{-1}(\Omega, \Sigma) : \int f^+ d\mu < \infty \text{ and } \int f^- d\mu < \infty\}. \qquad (1.10b)$$

The letter L stands for *Lebesgue space* (whose space-feature will be established later on) and L^1 means that this Lebesgue space is a subclass of the more general L^p family of spaces explored in [Dsh2].

Note that

$$\int |f| d\mu = \int f^+ d\mu + \int f^- d\mu, \qquad (1.10c)$$

which is due to $|f| = f^+ + f^-$ and Proposition 1.8. In light of (1.10c), expression (1.10b) can be rewritten as

$$L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$$

$$= \{f \in \mathcal{C}^{-1}(\Omega, \Sigma) : \int |f| d\mu < \infty\}. \qquad (1.10d)$$

If a measurable space (Ω, Σ) is specified, the notation $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ or $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ is sometimes shortened to $f \in \mathbb{L}(\mu)$ or $f \in L^1(\mu)$.

(ii) If $\Omega = \mathbb{R}^n$, $\Sigma = \mathcal{B}$, and μ is the Borel - Lebesgue measure λ and if the integral of the function f in (1.10) exists, it is called the *Lebesgue integral of* f. If f is λ-integrable, we write $f \in L^1(\lambda)$.

(iii) If $\Omega = \mathbb{R}$, $\Sigma = \mathcal{B}$, and $\mu = \mu_F$ (a Borel - Lebesgue - Stieltjes measure induced by an extended distribution function F), and if $g \in \mathbb{L}(\Omega, \Sigma, \mu_F; \overline{\mathbb{R}})$, then the integral in (1.10) is called the *Lebesgue - Stieltjes integral of* g; and we write $g \in L^1(\mu_F)$ if g is μ_F-integrable.

1.11 Remark. Let $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ be the space of all extended real-valued random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$. From

Example 4.5(i), Chapter 5, we recall that any measurable function $X \in \mathcal{C}^{-1}(\Omega, \Sigma)$ on $(\Omega, \Sigma, \mathbb{P})$ is a r.v.. If $X \in L^1(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}})$, then the numeric value $\int X d\mathbb{P}$ is called the *expectation of the r.v.* X, in notation, $\mathbb{E}X$ or $\mathbb{E}[X]$. From the definition, $\mathbb{E}X$ makes sense only if X is $\mathbb{P}$-integrable, which is equivalent to $\int |X| d\mathbb{P} < \infty$ (often perceived as an ambiguous restriction in probability books). □

1.12 Proposition. *The integral is a linear*, *monotone*, *non-decreasing functional on the vector space* $L^1(\Omega, \Sigma, \mu)$. □

(See Problem 1.8.)

1.13 Proposition. *For each function* $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, *the following holds*:

$$|\int f d\mu| \leq \int |f| d\mu \,.$$

Proof. Obviously, $|f| \geq f$ and $|f| \geq -f$. Thus, by Proposition 1.12, we have

$$\int |f| d\mu \geq \int f d\mu, \ \int |f| d\mu \geq \int (-f) d\mu = -\int f d\mu$$

and consequently,

$$\int |f| d\mu \geq |\int f d\mu|. \qquad \square$$

1.14 Notation. Let $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ and $A \in \Sigma$. Then, denote

$$\int\limits_A f d\mu = \int f \mathbf{1}_A \, d\mu.$$

Specifically, it follows that

$$\int\limits_\Omega f d\mu \ = \int f d\mu. \qquad \square$$

Recall Definitions and Remarks 5.7, Chapter 5, regarding properties that hold almost everywhere and the quotient sets of functions induced by a measure.

1.15 Lemma. *Let (Ω, Σ, μ) be a measure space and let $f \in \mathcal{C}_+^{-1}$. Then $\int f d\mu = 0$ if and only if $f \in [\mathbf{0}]_\mu^+$, where $[\mathbf{0}]_\mu^+ = [\mathbf{0}]_\mu \cap \mathcal{C}_+^{-1}$.*

Proof. Denote $N = \{f > 0\}$ (which is an element of Σ).

(i) Let $f \in [\mathbf{0}]_\mu^+$. Then $N \in \mathcal{N}_\mu \subseteq \Sigma$. Let $s_n = n\mathbf{1}_N$ $(\in \Psi_+)$, $n = 1, 2, \ldots$. Therefore,

$$\int s_n d\mu = n\mu(N) = 0, \quad \text{for all } n.$$

Denote $s = \sup\{s_n\}$. Then, by Proposition 5.6 (iv), Chapter 5, $s \in \mathcal{C}_+^{-1}$ and

$$\int s d\mu = \sup\{\int s_n d\mu\} = 0.$$

Furthermore, $f = s_n = 0$ on N^c. Although f is arbitrary on N and, in particular, not necessarily ∞, we have that $s_n \uparrow \infty$ pointwise on N. Consequently, $f \leq s$ on Ω which, by monotonicity (Proposition 1.8), yields

$$0 \leq \int f d\mu \leq \int s d\mu = 0$$

and hence $\int f d\mu = 0$.

(ii) Now let $\int f d\mu = 0$. Denote

$$N_n = \{f \geq \tfrac{1}{n}\} \, (= f^*\big([\tfrac{1}{n}, \infty]\big)), \; n = 1, 2, \ldots \, .$$

Obviously, $N_n \in \Sigma$ and $N_n \uparrow N$, where

$$N = \bigcup_{n=1}^{\infty} N_n = \{f > 0\} \in \Sigma.$$

By continuity from below of μ,

$$\lim_{n \to \infty} \mu(N_n) = \mu(N). \tag{1.15}$$

Clearly, $nf \geq \mathbf{1}_{N_n}$. Again, by monotonicity (Proposition 1.8), we have that

$$0 = \int f d\mu \geq \int \tfrac{1}{n}\, \mathbf{1}_{N_n} d\mu = \tfrac{1}{n}\, \mu(N_n)\,,$$

which leads to $\mu(N_n) = 0,\ n = 1, 2, \ldots$. From (1.15) it follows that $\mu(N) = 0$ and hence $N \in \mathcal{N}_\mu$. Therefore, $f \in [\mathbf{0}]^+_\mu$. □

1.16 Proposition. *Let* (Ω, Σ, μ) *be a measure space and let* $f, g \in \mathcal{C}^{-1}_+$ *such that* $f = g \pmod{\mu}$. *Then*

$$\int f d\mu = \int g d\mu.$$

Proof. By Problem 5.2(iii), Chapter 5, we have that $N = \{f \neq g\} \in \Sigma$. Therefore, by the above assumption regarding f and g, $N \in \mathcal{N}_\mu$ and the functions $f\mathbf{1}_N$ and $g\mathbf{1}_N$ are elements of the quotient set $[\mathbf{0}]^+_\mu$. By Lemma 1.15, it follows that

$$\int f\mathbf{1}_N d\mu = \int g\mathbf{1}_N d\mu = 0.$$

On the other hand,

$$\begin{aligned}\int f d\mu &= \int (f\mathbf{1}_N + f\mathbf{1}_{N^c}) d\mu \\ &= \int f\mathbf{1}_N d\mu + \int f\mathbf{1}_{N^c} d\mu = \int f\mathbf{1}_{N^c} d\mu.\end{aligned}$$

Similarly,

$$\int g d\mu = \int g\mathbf{1}_{N^c} d\mu.$$

The statement follows from $f\mathbf{1}_{N^c} = g\mathbf{1}_{N^c},\ \forall \omega \in \Omega$. Indeed,

$$\int f = \int\limits_N f + \int\limits_{N^c} f = \int\limits_{N^c} f \ \text{ and } \ \int g = \int\limits_N g + \int\limits_{N^c} g = \int\limits_{N^c} g$$

with $f = g$ on N^c. □

1.17 Proposition. *Let* (Ω, Σ, μ) *be a measure space and let* $f, g \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ *such that* $|f| \leq g$ *a.e. Then* $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ *implies that* $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$.

Proof. Let $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. Then by Proposition 6.3(iii), Chapter 5, we have that

$$g' = \sup\{g, |f|\} \in \mathcal{C}^{-1}.$$

Clearly,

$$|f| \leq g' \text{ everywhere and } g' = g \,(\text{mod } \mu) \;\; (\text{show it}),$$

and by Problem 1.12, $g' \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. Then, by Problem 1.9, $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. □

1.18 Proposition. *Let* $f, g \in \mathcal{C}^{-1}(\Omega, \Sigma)$ *and* f *or* $g \in L^1(\Omega, \Sigma, \mu)$. *Then*

$$\int_A f d\mu = \int_A g d\mu, \;\; \textit{for each } A \in \Sigma, \tag{1.18}$$

implies that $f = g$ (mod μ). □

(See Problem 1.22.)

Theorem 1.19 and Corollary 1.20 modify and, to some extent, refine Proposition 1.18.

1.19 Theorem. *If* μ *is* σ*-finite,* $f, g \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, *and*

$$\int_A f d\mu \leq \int_A g d\mu, \textit{ for each } A \in \Sigma, \tag{1.19}$$

then $f \leq g$ μ-a.e. *on* Ω.

Proof.

a) Let μ be finite. Denote

$$A_n = \{f \geq g + \tfrac{1}{n}, \; |g| \leq n\}$$

$$= \{f \geq g + \tfrac{1}{n}\} \cap \{|g| \leq n\}.$$

Then, since by our assumption, $\int\limits_A f d\mu \leq \int\limits_A g\, d\mu$ for each $A \in \Sigma$, we have

$$\begin{aligned} L: &= \int\limits_{A_n} g d\mu \geq \int\limits_{A_n} f d\mu \geq \int\limits_{A_n} (g + \tfrac{1}{n}) d\mu \\ &= M: = \int\limits_{A_n} g d\mu + \tfrac{1}{n}\mu(A_n). \end{aligned} \tag{1.19a}$$

On the other hand, by Proposition 1.13,

$$\left| \int\limits_{A_n} g d\mu \right| \leq \int\limits_{A_n} |g| d\mu \leq \int\limits_{A_n} n d\mu = n\mu(A_n) < \infty.$$

Therefore, from (1.19a) and because $\int\limits_{A_n} g\, d\mu$ is finite, $L \geq M$, which yields that $\mu(A_n) = 0$, for each n.

Thus,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \textstyle\sum_{n=1}^{\infty} \mu(A_n) = 0.$$

On the other hand, from

$$\bigcup_{n=1}^{\infty} A_n = \underbrace{\left(\bigcup_{n=1}^{\infty} \{f \geq g + \tfrac{1}{n}\}\right)}_{\{f > g\}} \cap \underbrace{\left(\bigcup_{n=1}^{\infty} \{\,|\, g \,|\, \leq n\}\right)}_{\{\omega : g(\omega) \text{ is finite}\}}$$

we conclude that $\mu\{f > g\text{: } g \text{ is finite}\} = 0$. Hence,

$$\mu\{f > g : g \geq -\infty\} = 0.$$

Letting $B_n = \{g = -\infty, f \geq -n\}$ we have

$$-\infty \cdot \mu(B_n) = \int\limits_{B_n} g d\mu \geq \int\limits_{B_n} f d\mu \geq -n\mu(B_n)$$

and therefore,

$$-n\mu(B_n) \leq -\infty\mu(B_n)$$

or, equivalently, $n\mu(B_n) \geq \infty\mu(B_n)$. This holds true if and only if $\mu(B_n) = 0$ (as the consequence of the agreement that $\infty \cdot 0 = 0$). Thus,

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \mu\{f > g,\, g = -\infty\} = 0.$$

In summary, we proved that $\mu\{f > g\} = 0$ implies that $f \leq g$ μ-a.e. on Ω.

b) Now, let μ be σ-finite and let $\mu_n = \mathrm{Res}_{\Sigma\cap\Omega_n}\mu$. Then

$$\int\limits_{A\cap\Omega_n} f d\mu = \int\limits_{A} \mathbf{1}_{\Omega_n} f d\mu \leq \int\limits_{A} \mathbf{1}_{\Omega_n} g d\mu$$

and hence $f \leq g$ μ-a.e. on Ω_n. The rest of this case is obvious. □

The reader can easily conclude the following.

1.20 Corollary. *If μ is σ-finite, $f, g \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, and*

$$\int\limits_{A} f d\mu = \int\limits_{A} g d\mu, \text{ for each } A \in \Sigma, \tag{1.20}$$

then $f = g$ μ-a.e. on Ω. □

(For a related discussion also see Problem 1.23.)

We conclude this section with a proposition below that is often cited in the sequel and have the reader prove it in Problem 1.14.

1.21 Proposition. *Each function $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ is finite μ-a.e. on Ω.* □

PROBLEMS

1.1 Prove Proposition 1.3.

(*i*) *For each measurable set $A \in \Sigma$,*

$$\int \mathbf{1}_A \, d\mu = \mu(A).$$

(ii) *The integral* $\int$ *is a nonnegative linear functional, i.e.,*

$$\int (as + bt)\, d\mu = a \int s d\mu + b \int t d\mu,$$

where $s, t \in \Psi_+$ *and* $a, b \in \mathbb{R}_+$.

(iii) *For any two nonnegative simple functions,* $s,\ t \in \Psi_+$, *such that* $s \leq t$, *it holds that*

$$\int s d\mu \leq \int t d\mu \ (\textit{monotonicity}).$$

1.2 Prove Corollary 1.6. *For* $\{s_n\}\uparrow,\ \{t_n\}\uparrow \subseteq \Psi_+$ *such that* $\sup\{s_n\} = \sup\{t_n\}$ *it holds that*

$$\sup\{\int s_n d\mu\} = \sup\{\int t_n d\mu\}.$$

[*Hint*: Use the fact that $s_j \leq \sup\{t_n\}$ and $t_k \leq \sup\{s_n\}$.]

1.3 Let (Ω, Σ, μ) be a measure space and $c > 0$. Show that for each $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$,

$$\int f d(c\mu) = c \int f d\mu.$$

1.4 Let $f \in \mathcal{C}^{-1}(\Omega, \Sigma, \mu; \mathbb{R})$ such that $\mu = \sum_{i=0}^{n} c_i \varepsilon_{a_i}$ (where $c_i \geq 0,\ i = 0, \ldots, n$). Show that $\int f d\mu$ exists and equals

$$\int f d\mu = \sum_{i=0}^{n} c_i f(a_i).$$

1.5 Let $f \in \mathcal{C}^{-1}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ and suppose that there are two nonnegative integrable functions u and v such that $f = u - v$. Show that $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ and that

$$\int f d\mu = \int u d\mu - \int v d\mu.$$

1.6 Show that $L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ is a vector space over $\mathbb{R}$.

1.7 Show that $(L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}), \mathbb{R}, +, \cdot, \preceq)$ is a Riesz space.

1.8 Prove Proposition 1.12. *The integral is a linear, monotone nondecreasing functional on the space* $L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$.

1.9 Let (Ω, Σ, μ) be a measure space and let $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Show that $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ if and only if there exists $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ such that $|f| \leq g$.

1.10 Show that $\{L^1(\Omega, \Sigma, \varepsilon_a; \overline{\mathbb{R}}) : a \in \Omega\} = \{L^1(\Omega, \Sigma, \varepsilon_a; \mathbb{R}): a \in \Omega\}$.

1.11 Show that if $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ then $f \in [\mathbf{0}]_\mu$ yields that $\int f d\mu = 0$. Does the converse hold true?

1.12 Let (Ω, Σ, μ) be a measure space and let $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, $g \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ such that $f - g \in [\mathbf{0}]_\mu$. Show that $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ and that $\int f d\mu = \int g d\mu$.

1.13 Generalize Proposition 1.16. *Let* (Ω, Σ, μ) *be a measure space and let* $f, g \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ *such that* $f = g$ *(mod* μ*) and* $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. *Then*

$$\int f d\mu = \int g d\mu$$

and $g \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$.

1.14 Show that each function $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ is finite μ-a.e. on Ω. [*Hint*: Let $A = \{|f| = \infty\}$. Show that $a\mu(A) < \infty$, $\forall a \in \mathbb{R}_+$, and then show that the inequality $\lim_{n \to \infty} n\mu(A) < \infty$ implies that $\mu(A) = 0$.]

1.15 Show that for $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$, $\int_A f d\mu = 0$ for each $A \in \Sigma$ if and only if $f \in [\mathbf{0}]_\mu$. Does this assertion generalize Lemma 1.15?

1.16 Show that

$$L^1(\Omega, \Sigma, \mu_1 + \mu_2; \overline{\mathbb{R}}) = L^1(\Omega, \Sigma, \mu_1; \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma, \mu_2; \overline{\mathbb{R}}).$$

1.17 Let π_α be a Poisson measure and let $f \in \mathcal{C}_+^{-1}(\mathbb{R}, \mathcal{B}; \overline{\mathbb{R}})$. Show that

$$\int f d\pi_\alpha = \sum_{n=0}^{\infty} e^{-\alpha} \frac{\alpha^n}{n!} f(n).$$

1.18 Under the conditions of Problem 1.5 assume that

a) $f(x) = x^2$,

b) $f(x) = \Gamma(x+1)\, \mathbf{1}_{(0,\infty)}(x)$, where $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ is the gamma function,

and find in each case the integral of f relative to measure π_α.

1.19 Let $\{\mu_n\}$ be a sequence of measures on a measurable space (Ω, Σ), $\{c_n\}$ be a sequence of positive real numbers, and let $\mu = \sum_{n=1}^{\infty} c_n \mu_n$ [which is a measure on (Ω, Σ)]. Show that for every $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$,

$$\int f d\mu = \sum_{n=1}^{\infty} c_n \int f d\mu_n.$$

1.20 Let μ and ν be two measures on (Ω, Σ) such that $\mu \leq \nu$. Show that for each $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}})$, the integral $\int f d(\nu - \mu)$ makes sense, $f \in L^1(\Omega, \Sigma, \nu - \mu; \overline{\mathbb{R}})$, and that

$$\int f d(\nu - \mu) = \int f d\nu - \int f d\mu.$$

1.21 Let μ and ν be two measures on (Ω, Σ) such that $\mu \leq \nu$. Show that for each $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$, $\int f d\mu \leq \int f d\nu$.

1.22 Prove Proposition 1.18. *Show that if* $f, g \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ *and* f *or* $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, *then*

$$\int_A f d\mu = \int_A g d\mu \ \textit{for each } A \in \Sigma \tag{P1.22}$$

yields that $f = g$ *(mod* μ*).*

1.23 Show by a counterexample that dropping the condition f or $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ in Problem 1.22 need not yield $f = g$ (mod μ) even if f and g are nonnegative.

1.24 Is $\mathbb{L}(\Omega, \Sigma; \mu, \overline{\mathbb{R}})$ a linear space?

1.25 Let $g \in \mathcal{C}^{-1}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ such that $\int_A g\, d\mu \geq 0$ for all $A \in \Sigma$. Show that $g \geq 0$ μ-a.e.

1.26 In Problem 1.15, it was stated that $f \in [\mathbf{0}]_\mu$ if and only if $\int_A f d\mu = 0$. Corollary 1.20 stated that $\int_A f d\mu = \int_A g d\mu$ if and only if $f = g$ modulo μ, provided that at least one of the functions f or $g \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$. These two seem to be similar if in Problem 1.15 we replace f with $f - g$. Is it so?

1.27 Is $L^1(\Omega, \Sigma; \mu, \overline{\mathbb{R}}_+)$ a proper subset of $\mathcal{C}_+^{-1}$? What is the relationship between $\mathbb{L}(\Omega, \Sigma; \mu, \overline{\mathbb{R}}_+)$ and $\mathcal{C}_+^{-1}$?

1.28 Let $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Show that $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ if and only if for each $\varepsilon > 0$, there is a function $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+)$ such that

$$\int_{\{|f| \geq g\}} |f| d\mu \leq \varepsilon.$$

1.29 Show by a counterexample that L^1 is not an $\mathcal{X}$-space.

NEW TERMS:

Lebesgue, Henri, 361
Fatou, Pierre, 361
Fubini, Guido, 361
Riesz, Frigyes, 361
Nikodym, Otto, 361
Radon, Johann, 361
integral of a nonnegative simple function, 362
Dirichlet function, 364
integral of an extended nonnegative function, 365
integral with respect to Dirac point mass, 366
moment generating function, 367
integral of an extended real-valued function, 369
μ-integrable function, 370
$L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, Lebesgue space, 370
$\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$-space 370
Lebesgue integral, 370
Lebesgue - Stieltjes integral, 370
expectation of a random variable, 371
L^1 as a vector space, 371
integral as a linear functional on L^1, 371
L^1 as a vector lattice, 378

2. MAIN CONVERGENCE THEOREMS

The following result is one of the basic convergence theorems a special case of which (Corollary 2.2) was originally proved by the Italian mathematician, Beppo Levi (1875 – 1961).

2.1 Theorem (of Monotone Convergence). *Let* $\{f_n\}\uparrow \subseteq \mathcal{C}_+^{-1}$. *Then*

$$\int \sup\{f_n\} d\mu = \sup\{\int f_n d\mu\}. \tag{2.1}$$

Proof. Let $f = \sup\{f_n\}$. Then, by Proposition 5.6(iv), Chapter 5, $\sup\{f_n\} \in \mathcal{C}_+^{-1}$. Thus, the integral on the left-hand side of (2.1) makes sense. On the other hand, for each element $f_n \in \mathcal{C}_+^{-1}$, there exists a monotone nondecreasing sequence of nonnegative simple functions $\{s_{k_n}^{(n)}\}\uparrow \subseteq \Psi_+$ such that $\sup\{s_{k_n}^{(n)}\colon\ k_n = 1, 2, \dots\} = f_n$. Let

$$t_k = \max\{s_k^{(1)}, \dots, s_k^{(k)}\}.$$

Because Ψ_+ is a lattice, it follows that $t_k \in \Psi_+,\ k = 1, 2, \dots$. Furthermore, $\{t_k\}$ is monotone nondecreasing. Because $\{f_n\}$ is monotone nondecreasing, we have

$$s_k^{(1)} \le f_1 \le f_k,\ \ s_k^{(2)} \le f_2 \le f_k, \dots,\ s_k^{(k)} \le f_k\ ,$$

and hence $s_k^{(i)} \le f_k,\ i = 1, \dots, k,$ which leads to

$$t_k \le f_k \tag{2.1a}$$

and

$$\sup\{t_k\} \le \sup\{f_k\} = f. \tag{2.1b}$$

On the other hand, $t_{k_n} \ge s_{k_n}^{(n)}$ for $k_n \ge n;$ this yields

$$\sup\{s_{k_n}^{(n)}\colon\ k_n = 1, 2, \dots\} = f_n \le \sup\{t_k\colon\ k = 1, 2, \dots\},$$

and, consequently,

$$\sup\{f_n\} = f \leq \sup\{t_k\}. \tag{2.1c}$$

Thus, by (2.1b) and (2.1c),

$$f = \sup\{t_k\}\,.$$

Now, the facts that $f = \sup\{t_k\}$ and that $\{t_k\}$ is monotone non-decreasing imply that

$$\int f d\mu = \sup \int t_k d\mu.$$

In as much as $t_k \leq f_k$ by (2.1a), we have by Proposition 1.8

$$\int t_k d\mu \leq \int f_k d\mu\,,$$

which yields

$$\int f d\mu = \sup \int t_k d\mu \leq \sup \int f_k d\mu.$$

Finally, the inverse inequality holds due to $f_n \leq f$ and Proposition 1.8. □

2.2 Corollary (Beppo Levy). *Let* $\{f_n\} \subseteq \mathcal{C}_+^{-1}$. *Then*

$$\int \left[\sum_{n=1}^{\infty} f_n\right] d\mu = \sum_{n=1}^{\infty} \int f_n d\mu. \qquad \square$$

(See Problem 2.1.)

The monotone convergence theorem can be generalized for an arbitrary monotone sequence under a minor constraint.

2.3 Theorem (Generalized Monotone Convergence Theorem). *Let* $\{f_n\}\uparrow\ \subseteq \mathcal{C}^{-1}(\Omega, \Sigma; \bar{\mathbb{R}})$ *and* $g \in \mathcal{C}^{-1}(\Omega, \Sigma; \bar{\mathbb{R}})$ *such that* $f_n \geq g$ *for all* n *and* $\int g d\mu > -\infty$. *Then*,

$$\sup\{\int f_n\, d\mu\} = \int \sup\{f_n\}\, d\mu\,. \qquad \square$$

(See Problem 2.2.)

The next is the very often-used Fatou's lemma.

2.4 Lemma (Fatou). *Let* (Ω, Σ, μ) *be a measure space and let* $\{f_n\} \subseteq \mathcal{C}_+^{-1}(\Omega, \Sigma)$. *Then*

$$\int \underline{\lim}\, f_n d\mu \leq \underline{\lim} \int f_n d\mu.$$

Proof. By Proposition 5.6 (v), Chapter 5, $\underline{\lim}\ f_n \in \mathcal{C}_+^{-1}$ and by Proposition 5.6 (iv), Chapter 5,

$$g_n = \inf\{f_k \colon k \geq n\} \in \mathcal{C}_+^{-1}.$$

Clearly, the sequence $\{g_n\}$ is monotone nondecreasing and hence

$$\sup\{g_n\} = \underline{\lim}\, f_n \ \text{ and } \ g_n \leq f_k \text{ , for all } k \geq n.$$

By monotonicity of the integral,

$$\int g_n d\mu \leq \int f_k d\mu\,,\ \forall k \geq n,\ n = 1, 2, \dots\,,$$

which implies that

$$\int g_n d\mu \leq \inf\{\int f_k d\mu \colon k \geq n\}.$$

Finally, by the monotone convergence theorem,

$$\int \sup\{g_n\} d\mu = \sup\{\int g_n d\mu\}$$

$$\leq \sup \inf\{\int f_k d\mu\} = \underline{\lim} \int f_n d\mu. \qquad \square$$

2.5 Definition. Let $f,\ \{f_n\} \subseteq L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. The sequence $\{f_n\}$ is said to *converge to* f *in mean* if

$$\int \mid f - f_n \mid d\mu \ \to\ 0. \qquad \square$$

We now formulate and prove one of the central results in the theory of integration. As with the monotone convergence theorem,

the following theorem enables us to interchange the limit and the integral for a pointwise convergent sequence of functions. However, it does not require that the sequence be monotone nondecreasing and nonnegative. On the other hand, the sequence needs to have an integrable dominating function, and thus it is not a generalization of the monotone convergence theorem.

2.6 Theorem (Lebesgue's Dominated Convergence). *Let* (Ω, Σ, μ) *be a measure space and let* $\{f_n\} \subseteq \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ *be a pointwise a.e. convergent sequence. Suppose that there is a* μ*-integrable function* $g\ (\in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}))$ *such that* $g \geq 0$, *and that* $|f_n| \leq g,\ n = 1, 2, \ldots$. *Then the following are true.*

(i) *There exists at least one function* $f \in \mathcal{C}^{-1}$, *such that* $f < \infty$, *to which the sequence* $\{f_n\}$ *converges a.e. in the topology of pointwise convergence.*

(ii) $f \in L^1(\Omega, \Sigma, \mu; \mathbb{R})$ *and* $\{f_n\} \subseteq L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$.

(iii) *The sequence* $\{f_n\}$ *converges to* f *in mean, that is,*

$$\lim_{n\to\infty} \int \mid f - f_n \mid d\mu = 0.$$

(iv) $\int f d\mu = \lim_{n\to\infty} \int f_n d\mu$.

Proof.

(i) By our assumption, there is a negligible set Π such that

$$\lim_{n\to\infty} f_n(\omega)$$

exists for all $\omega \in \Pi^c$ and there is a μ-null set $N_1 \supseteq \Pi$. Therefore, $N_1^c \subseteq \Pi^c$ and

$$\lim_{n\to\infty} f_n(\omega)$$

exists for all $\omega \in N_1^c$. Because $g \in L^1(\Omega, \Sigma, \mu)$, by Proposition 1.21, it follows that g is finite μ-a.e. on Ω, that is, there is a μ-null set N_2 such that $g(\omega) < \infty$ for all $\omega \in N_2^c$. Define the function

$$f(\omega) = \mathbf{1}_A(\omega) \lim_{n\to\infty} f_n(\omega), \tag{2.6}$$

where $A = (N_1 \cup N_2)^c$. Clearly, f_n converges to f pointwise μ-a.e. on Ω and hence, by Proposition 5.6(iii) and (vi), Chapter 5, $f \in \mathcal{C}^{-1}$. Indeed, because f_n and $\mathbf{1}_A \in \mathcal{C}^{-1}$, it follows that $f_n \mathbf{1}_A \in \mathcal{C}^{-1}$ and that $f_n \mathbf{1}_A \to f$ in the topology of pointwise convergence; the latter implies that $f_n \to f$ pointwise μ-a.e. on Ω.

(ii) From (2.6) it follows that on set A, $\lim_{n\to\infty} f_n = f$; in addition, $\{f_n\}$ is dominated by a finite function g on A. Thus, $|f| \le g$ on A and, due to (2.6), $f = 0$ on A^c. Hence,

$$|f| < \infty \text{ and } |f| \le g, \ \forall \omega \in \Omega.$$

By Proposition 1.17 and in as much as $|f| < \infty$, $f \in L^1(\Omega, \Sigma, \mu)$. Also by Proposition 1.17, $\{f_n\} \subseteq L^1(\Omega, \Sigma, \mu)$.

(iii) We prove that f_n is convergent in mean to f, that is,

$$\int \mid f - f_n \mid d\mu \to 0.$$

Let $g_n = |f - f_n|$ $(\in \mathcal{C}_+^{-1}(\Omega, \Sigma))$. Then, $0 \le g_n \le |f| + g$. Because

$$|f| + g \in L^1(\Omega, \Sigma, \mu),$$

it follows that $g_n \in L^1(\Omega, \Sigma, \mu)$, by Problem 1.9.

[Observe that because linearity of the integral holds just on L^1, we do need to show that $g_n \in L^1$ which would lead to

$$\int |f| + g - g_n = \int (|f| + g) + \int (-g_n).]$$

Applying Fatou's lemma to the sequence $\{|f| + g - g_n\}$, we have:

$$\int \underline{\lim}(|f| + g - g_n) d\mu \le \underline{\lim} \int (|f| + g - g_n) d\mu$$

$$= \int (|f| + g) d\mu + \underline{\lim} \int (-g_n) d\mu$$

$$= \int (|f| + g) d\mu - \overline{\lim} \int g_n d\mu \tag{2.6a}$$

Because $f_n \to f$ a.e., then $g_n \to 0$ a.e., and hence $|f| + g - g_n \to |f| + g$ a.e. which implies that

$$\underline{\lim}\,(|f| + g - g_n) = |f| + g \text{ a.e.}$$

By Proposition 1.16,

$$\int \underline{\lim}\,(|f| + g - g_n)d\mu = \int (|f| + g)d\mu,$$

which, together with inequality (2.6a), yields

$$\int (|f| + g)d\mu \leq \int (|f| + g)d\mu - \overline{\lim} \int g_n d\mu$$

or, equivalently,

$$\overline{\lim} \int g_n d\mu \leq 0. \tag{2.6b}$$

Because $g_n \geq 0$, (2.6b) reveals that

$$\overline{\lim} \int g_n d\mu = 0$$

and thus

$$\lim \int |f - f_n| d\mu = 0,$$

which proves (iii).

Now (iv) is due to Problem 2.6. □

2.7 Examples.

(i) We evaluate $\lim\limits_{n\to\infty} \int_0^1 nx(1-x)^n \lambda(dx)$. First observe that the sequence $\{nx(1-x)^n\}$ is convergent to the function **0** pointwise on $[0, 1]$. However, it is an easy exercise to show that the sequence $\{nx(1-x)^n\}$ does not converge to **0** uniformly. Otherwise, we could interchange the limit and the integral. (See Problem 3.12 of the next section.) Fortunately, the functions $nx(1-x)^n$ are uniformly bounded by 1. Therefore, function **1** can be taken as a pertinent integrable majorant function in the Lebesgue dominated convergence theorem. This enables us to interchange the limit and the integral and conclude that

$$\lim_{n\to\infty} \int_0^1 nx(1-x)^n \lambda(dx) = 0.$$

[We can verify this result by direct computation of the integral

$$\int_0^1 nx(1-x)^n dx = \tfrac{n}{(n+1)(n+2)}$$

and then passing to the limit.]

(ii) Calculate $\lim_{n\to\infty} \int_0^n \left(1+\frac{x}{n}\right)^n e^{-2x} \lambda(dx)$. Clearly,

$$\left(1+\tfrac{x}{n}\right)^n \mathbf{1}_{[0,n]}(x) e^{-2x} < e^{-x} \in L^1.$$

Hence, by the Lebesgue dominated convergence theorem,

$$\lim_{n\to\infty} \int \left(1+\tfrac{x}{n}\right)^n \mathbf{1}_{[0,n]}(x) e^{-2x} \lambda(dx)$$

$$= \int \lim_{n\to\infty} \left(1+\tfrac{x}{n}\right)^n \mathbf{1}_{[0,n]}(x) e^{-2x} \lambda(dx) = \int_0^\infty e^{-x} \lambda(dx) = 1. \qquad \square$$

2.8 Remark. Note that we treated $\int_0^1 nx(1-x)^n \lambda(dx)$ in Example 2.7(i) informally both as Lebesgue (L) and Riemann (R) integrals (in as much as they are identical in this case), although a formal relationship between the two needs to be rigorously developed, which we explore in Section 3. The same applies to Example 2.7(ii). In Problems 2.9 - 2.11 we therefore assume that the Lebesgue integrals are equal to Riemann integrals. $\square$

The following are analogues of the main convergence theorems (monotone convergence theorem, Fatou's Lemma, and Lebesgue's dominated convergence theorem) for measures, that are often used in probability and control theory. The theorems are essentially based on the results of Onésimo Hernández-Lerma and Jean Lasserre [Her], which were established in 2000 under weaker conditions than those in previous texts and papers.

2.9 Lemma (Fatou). *Let $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$ and $\{\mu, \mu_1, \mu_2, \dots\}$ be a sequence of measures on Σ such that for each $A \in \Sigma$, $\underline{\lim}\, \mu_n(A) \geq \mu(A)$. Then*

$$\int f d\mu \leq \underline{\lim} \int f d\mu_n.$$

Proof. Let $\{s_k\} \uparrow \ \subseteq \Psi_+(\Omega, \Sigma)$ such that $s_k \uparrow \ f$ and

$$s_k = \sum_{j=1}^{m_k} \alpha_j^{(k)} \mathbf{1}_{A_{jk}}.$$

Hence, for each $k = 1, 2, \ldots,$

$$\underline{\lim} \int f d\mu_n \geq \underline{\lim}_{n\to\infty} \int s_k d\mu_n = \underline{\lim}_{n\to\infty} \sum_{j=1}^{m_k} \alpha_j^{(k)} \mu_n(A_{j_k})$$

$$\sum_{j=1}^{m_k} \alpha_j^{(k)} \underline{\lim}_{n\to\infty} \mu_n(A_{j_k}) \geq \int s_k d\mu.$$

The statement now follows from the definition of integral. □

2.10 Theorem (Lebesgue's Dominated Convergence for Measures). *Let $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$ and $\{\nu, \mu, \mu_1, \mu_2, \ldots\}$ be a sequence of measures on Σ such that for each $A \in \Sigma$, $\mu_n(A) \to \mu(A)$, $\mu_n \leq \nu$, and $\int f d\nu < \infty$. Then*

$$\lim_{n \to \infty} \int f d\mu_n = \int f d\mu.$$

Proof. Because $\mu_n \leq \nu$, it is easy to verify that $\nu - \mu_n$ is a measure on Σ. Due to Problem 1.25,

$$\int f d(\nu - \mu_n) = \int f d\nu - \int f d\mu_n.$$

Furthermore,

$$\underline{\lim}(\nu - \mu_n)(A) = \nu(A) - \overline{\lim} \mu_n(A)$$

$$= \nu(A) - \lim_{n \to \infty} \mu_n(A) = \nu(A) - \mu(A) = (\nu - \mu)(A).$$

The last inequality holds true, because obviously $\mu \leq \nu$ and hence $\nu - \mu$ is a measure on Σ. Now, all conditions of Fatou's Lemma 2.9 are met for the sequences $\{\nu - \mu_n\}$ and $\{\mu_n\}$ and therefore,

$$\int f d(\nu - \mu) = \int f d\nu - \int f d\mu \leq \underline{\lim} \int f d(\nu - \mu_n)$$

$$= \underline{\lim}(\int f d\nu - \int f d\mu_n) = \int f d\nu - \overline{\lim} \int f d\mu_n$$

and

$$\int f d\mu \leq \underline{\lim} \int f d\mu_n.$$

Combining both inequalities we have

$$\overline{\lim} \int f d\mu_n \leq \int f d\mu \leq \underline{\lim} \int f d\mu_n$$

and hence, the statement. □

To prove the Theorem of Monotone Convergence for measures we need the notion of setwise convergence.

2.11 Definition. Let (Ω, Σ) be a measurable space and $\{\mu_n\}$ be a sequence of measures on (Ω, Σ). We say that $\{\mu_n\}$ converges to a set function μ *setwise* if $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ exists for each $A \in \Sigma$. The set function μ is called the *setwise limit* of $\{\mu_n\}$. □

2.12 Proposition. *The setwise limit μ of $\{\mu_n\}$ has the following properties*.

(i) *μ is monotone and additive.*

(ii) *Let $\{A_1, A_2, \dots\}$ be a sequence of pairwise disjoint sets from Σ and $A_n \subseteq A \in \Sigma$. Then*

$$\sum_{k=1}^{\infty} \mu(A_k) \leq \mu(A). \tag{2.12}$$

Proof.

(i) Is trivial.

(ii) It can be verified directly from the definition of the setwise limit by using monotonicity and additivity or just due to Proposition 1.3(ii), Chapter 5. □

We are wondering what condition imposed on a sequence $\{\mu_n\}$ makes its setwise limit a measure. For instance, if the sequence $\{\mu_n\}$ is monotone nonincreasing, then the limiting set function μ need not be σ-additive, as we learn from Problem 2.12. The following theorem provides us with some sufficient conditions.

2.13 Theorem. *Let a sequence* $\{\mu_n\}$ *of measures on a measure space* (Ω, Σ) *be convergent to a set function* μ *setwise. Then* μ *is a measure if one of the following conditions holds*.

(i) $\{\mu_n\}$ *is a monotone nondecreasing sequence.*

(ii) μ *is finite*.

Proof. Let $\{A_k\}$ be a sequence of pairwise disjoint measurable sets with A as its union.

(i) $\{\mu_k\}$ is monotone nondecreasing, for each $m = 1, 2, \dots$. Thus

$$\mu_m(A) = \sum_{k=1}^{\infty} \mu_m(A_k) \leq \sum_{k=1}^{\infty} \mu(A_k),$$

which, combined with (2.12), yields the statement.

(ii) Because μ is finite, by Theorem 1.7(ii), Chapter 5, if μ is not σ-finite (which we assume), it would not be Ø-continuous. In other words, there is a monotone nondecreasing sequence $\{A_k\} \downarrow$ Ø of measurable sets such that $\lim_{k\to\infty} \mu(A_k) = \quad \varepsilon > 0$. Let $a_1 = b_1 = 1$ and suppose a_j and b_j are positive integers defined for all $j \leq n$. Furthermore, let $a_{n+1} > a_n$ such that

$$\tfrac{7}{8}\varepsilon \leq \mu_{a_{n+1}}(A_{b_n}).$$

[If there is no such a_{n+1}, then it would surely contradict our assumption that $\lim_{k\to\infty} \mu(A_k) = \varepsilon > 0$.] Now, let $b_{n+1} > b_n$ such that

$$\tfrac{1}{8}\varepsilon \geq \mu_{a_{n+1}}(A_{b_{n+1}}).$$

(Such a b_{n+1} should exist, because $\mu_{a_{n+1}}$ is Ø-continuous.) For $B_n := A_{b_n} \backslash A_{b_{n+1}}$, we have that $\mu_{a_{n+1}}(B_n) \geq \frac{3\varepsilon}{4}$. Therefore for j being odd and $j > k \geq 1$,

$$\mu_{a_j}\left(\sum_{n \text{ even}: n \geq k} B_n\right) \geq \tfrac{3}{4}\varepsilon.$$

Then, for $k \geq 1$,

$$\mu\left(\sum_{n \text{ even}:n\geq k} B_n\right) \geq \tfrac{3}{4}\varepsilon.$$

We can easily verify that the last inequality holds true also for all odd values of n. Consequently, for all $k \geq 1$,

$$\mu(A_{b_k}) = \mu\Big(\textstyle\sum_{s=1}^{\infty} B_s\Big) \geq \tfrac{3}{2}\varepsilon.$$

The latter contradicts the assumption that $\lim_{k\to\infty}\mu(A_k) = \varepsilon > 0$. □

The next theorem is of monotone convergence for a sequence of measures.

2.14 Theorem (Monotone Convergence). *Let $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$ and $\{\mu_1, \mu_2, \dots\}$ be a monotone nondecreasing sequence of measures on a measure space (Ω, Σ). Then there is a measure μ on (Ω, Σ, μ) such that $\mu_n(A) \to \mu(A)$ for all A of Σ and*

$$\lim_{n\to\infty}\int f d\mu_n = \int f d\mu. \tag{2.14}$$

Proof. Because $\{\mu_n\}$ is monotone nondecreasing, by Theorem 2.13 (i), the setwise limit μ of $\{\mu_n\}$ exists and is a measure on (Ω, Σ). In as much as f is nonnegative and $\mu_n \uparrow \mu$, the sequence $\{\int f d\mu_n\}$ is monotone nondecreasing and hence

$$\underline{\lim}\int f d\mu_n = \lim_{n\to\infty}\int f d\mu_n \leq \int f d\mu. \tag{2.14a}$$

The last inequality holds because of $\int f d\mu_n \leq \int f d\mu$, which, in turn, is due to Problem 1.26. On the other hand, from Fatou's Lemma 2.9 applied to our case,

$$\int f d\mu \leq \lim_{n\to\infty}\int f d\mu_n$$

which, combined with (2.14a), yields the statement. □

The convergence theorems below are for sequences of nonnegative functions and measures at once.

2.15 Lemma (Fatou for Measures and Nonnegative Functions). *Let* $\{\mu, \mu_1, \mu_2, \dots\}$ *be a sequence of measures on a measure space* (Ω, Σ) *and let* $\{f_n\} \subseteq \mathcal{C}_+^{-1}(\Omega, \Sigma)$ *such that for each* $A \in \Sigma$, $\underline{\lim}\, \mu_n(A) \geq \mu(A)$. *Then*

$$\int f d\mu \leq \underline{\lim} \int f_n d\mu_n, \tag{2.15}$$

where

$$f(\omega) := \underline{\lim} f_n(\omega), \ \ \omega \in \Omega. \tag{2.15a}$$

Proof. First assume that $\{f_n\} \subseteq \mathcal{C}_+^{-1}(\Omega, \Sigma)$. Then, for every fixed positive integer N and for every n,

$$\int f_n d\mu_n \geq \int \inf\{f_m \colon m \geq N\} d\mu_n. \tag{2.15b}$$

Applying the version of Fatou's Lemma 2.9 to the right-hand side of (2.15b) we have

$$\begin{aligned} \underline{\lim} \int f_n d\mu_n &\geq \underline{\lim} \int \inf\{f_m \colon m \geq N\} d\mu_n \\ &\geq \int \inf\{f_m \colon m \geq N\} d\mu. \end{aligned} \tag{2.15c}$$

Because $\{\inf\{f_m \colon m \geq N\}_N\} \uparrow \ f$ defined in (2.15a), applying the standard monotone convergence Theorem 2.1, we arrive at

$$\underline{\lim} \int f_n d\mu_n \geq \int \inf\{f_m \colon m \geq N\} d\mu = \int f d\mu. \qquad \square$$

The following generalization of Fatou's Lemma 2.15 is applied to arbitrary measurable functions $\{f_n\}$ and its proof is left to the reader. (Problem 2.13.)

2.16 Lemma (Fatou for Measures and Functions). *In the condition of Fatou's Lemma* 2.15, *let* $\{g, f_1, f_2, \dots\} \subseteq \mathcal{C}^{-1}(\Omega, \Sigma)$ *such that for all* n, $f_n \geq g$ *and*

$$\lim_{n \to \infty} \int g d\mu_n = \int g d\mu > -\infty.$$

Then,

$$\int f d\mu \leq \underline{\lim} \int f_n d\mu_n, \tag{2.16}$$

where

$$f(\omega) := \underline{\lim} f_n(\omega), \omega \in \Omega. \qquad \square$$

This is a further modification of the Lebesgue dominated convergence theorem for a sequence of measures and functions.

2.17 Theorem (Lebesgue's Dominated Convergence for Measures and Functions). *Let* $\{f_n\} \subseteq \mathcal{C}^{-1}(\Omega, \Sigma)$, $g \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$, *and* $\{\nu, \mu, \mu_1, \mu_2, \dots\}$ *be a sequence of measures on the measure space* (Ω, Σ) *such that*:

- (i) $\mu_n \leq \nu$.
- (ii) f_n converges to a function f in the topology of pointwise convergence.
- (iii) μ_n *converges to* μ *setwise*.
- (iv) $\int g d\nu < \infty$.
- (v) $|f_n| \leq g$.

Then,

$$\lim_{n\to\infty} \int f_n d\mu_n = \int f d\mu. \tag{2.17}$$

Proof. Consider Theorem 2.10 for which we use the conditions (i), (iii), and (iv). Then, applying Theorem 2.10 to g we have that

$$\lim_{n\to\infty} \int g d\mu_n = \int g d\mu.$$

Now, because $g \pm f_n \geq 0$ for all n, we have from Fatou's Lemma 2.16,

$$\underline{\lim} \int (g \pm f_n) d\mu_n \geq \int \underline{\lim} (g \pm f_n) d\mu.$$

On the other hand, because $\int g d\mu \leq \int g d\nu < \infty$,

$$\overline{\lim} \int f_n d\mu_n \leq \int f d\mu \leq \underline{\lim} \int f_n d\mu_n$$

which yields the assertion. $\square$

PROBLEMS

2.1 Prove Corollary 2.2.

2.2 Generalize the monotone convergence theorem: Let $\{f_n\}\uparrow \subseteq \mathcal{C}^{-1}(\Omega, \Sigma)$ and $g \in \mathcal{C}^{-1}(\Omega, \Sigma)$ such that $f_n \geq g$ $f$or all n and suppose that $\int g\, d\mu > -\infty$. Prove that

$$\sup\{\int f_n\, d\mu\} = \int \sup\{f_n\}\, d\mu\,.$$

2.3 Show that if $\int g d\mu = -\infty$, the Generalized monotone convergence theorem need not hold.

2.4 Let $\{f_n\}\downarrow \subseteq \mathcal{C}^{-1}$ and $g \in \mathcal{C}^{-1}$ such that $f_n \leq g$ for all n. If $\int g$ $d\mu < \infty$, show that

$$\inf\{\int f_n\, d\mu\} = \int \inf\{f_n\}\, d\mu.$$

2.5 Let (Ω, Σ, μ) be a measure space and let $\{A_n\} \subseteq \Sigma$. Prove that

$$\mu(\underline{\lim}\, A_n) \leq \underline{\lim}\, \mu(A_n),$$

and if $\mu < \infty$ that

$$\mu(\overline{\lim}\, A_n) \geq \overline{\lim}\, \mu(A_n).$$

2.6 Show that if $f_n \to f$ in mean then

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$$

2.7 Generalize Fatou's Lemma 2.4 in the following way. Let $\{f_n\} \subseteq \mathcal{C}^{-1}(\Omega,\Sigma)$ and $g \in \mathcal{C}^{-1}(\Omega,\Sigma)$ such that $g \leq f_n$ for all n. Let $\int g^- d\mu < \infty$. Show that

$$\int \underline{\lim}\, f_n\, d\mu \leq \underline{\lim} \int f_n\, d\mu.$$

2.8 Let $\{f_n\} \subset \mathcal{C}^{-1}(\Omega,\Sigma)$ and $g \in \mathcal{C}^{-1}(\Omega,\Sigma)$ such that $f_n \leq g$ for all n. Let $\int g^+ d\mu < \infty$. Show that

$$\overline{\lim} \int f_n \, d\mu \leq \int \overline{\lim} \, f_n \, d\mu.$$

2.9 Let

$$f_n(x) = \begin{cases} n, & 0 < x < \frac{1}{n} \\ 0, & \frac{1}{n} \leq x < \infty. \end{cases}$$

Show that $f_n \to 0$ λ-a.e. in the topology of pointwise convergence. Explain why

$$\int \lim_{n\to\infty} f_n \, \lambda(dx) \neq \lim_{n\to\infty} \int f_n \, \lambda(dx).$$

2.10 Let

$$f_n(x) = \begin{cases} n^2 x, & 0 \leq x \leq \frac{1}{n} \\ -n^2\left(x - \frac{2}{n}\right), & \frac{1}{n} \leq x \leq \frac{2}{n} \\ 0, & x \geq \frac{2}{n}. \end{cases}$$

Show that

$$\int \underline{\lim} \, f_n \, d\lambda < \underline{\lim} \int f_n \, d\lambda.$$

Explain why the Lebesgue dominated convergence theorem cannot be applied.

2.11 Use Lebesgue's dominated convergence Theorem 2.6 to prove that for all $\alpha > 0$,

$$\lim_{n\to\infty} \frac{n! n^\alpha}{\alpha(\alpha-1)\cdots(\alpha+n-1)} = \Gamma(\alpha), \tag{P2.11}$$

where $\Gamma(\alpha)$ is known to be the *gamma function* and is expressed as the improper Riemann integral

$$\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} dx. \tag{P2.11a}$$

2.12 Give an example of a monotone nonincreasing sequence of measures convergent to a set function μ setwise such that μ is not a measure.

2.13 Prove Fatou's Lemma 2.16.

2.14 Let $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ and suppose that $\{\Omega_n\}$ is a countable measurable partition of Ω. Show that

$$\int f d\mu = \sum_{n=1}^{\infty} \int_{\Omega_n} f d\mu.$$

NEW TERMS:

Levi, Beppo, 382
monotone convergence theorem for functions, 382
Beppo Levi's corollary, 383
monotone convergence theorem, generalized, 383
Fatou's lemma for functions, 384
convergence in mean, 384
Lebesgue's dominated convergence theorem for functions, 385
Hernández-Lerma, Onésimo, 388
Lasserre, Jean, 388
Fatou's lemma for measures, 388
Lebesgue's dominated convergence theorem for measures, 389
setwise convergence of measures, 390
setwise limit of measures 390
setwise convergence, criterion of, 391
monotone convergence theorem for measures, 392
Fatou's lemma for measures and nonnegative functions, 393
Fatou's lemma for measures and functions, 393
Lebesgue's dominated convergence theorem for measures and functions, 394
gamma function 396

3. LEBESGUE AND RIEMANN INTEGRALS ON $\mathbb{R}$

In this section we develop integration techniques in $L^1(\mathbb{R}, \mathcal{B}, \lambda; \overline{\mathbb{R}})$ (see Definition $1.10(ii)$). The key idea is to reduce the Lebesgue integral to the Riemann integral whenever it is possible in combination with the main convergence theorems. The Bernhard Riemann notion of an integral, which was a refinement from its inception by Augustin - Louis Cauchy in 1823, was introduced in 1854.

We begin with the concept of the Riemann integral of a bounded function on a compact interval suggested by the Frenchman Gaston (in some sources, Jean - Gaston) Darboux (1842 – 1917) in 1875. Although the construction below is self-contained, the reader may find it plausible to go over Section 9 (Example $9.9(vi)$), Chapter 3, of [Dsh1] for topological preliminaries of this construction.

Let $\Omega = [a, b]$ be a compact interval in $\mathbb{R}$. By Definition $1.7(ii)$, Chapter 1, the *partition* of $[a, b]$ is any ordered n-tuple $P = P(n) = P(a_0, \dots, a_n)$ with

$$P = \{a_0, \dots, a_n \in [a, b]\colon a = a_0 < a_1 < \cdots < a_n = b\}.$$

Let P_1 and P_2 be two partitions of $[a, b]$. We say P_2 is *finer* than P_1 if $P_1 \subseteq P_2$. P_2 is also said to be a *refinement* of P_1 (in notation $P_1 \preceq P_2$). Thus, if $\mathcal{P}$ is the set of all partitions on $[a, b]$, $\preceq$ is a partial order on $\mathcal{P}$.

Denote by $\mathcal{C}_b^{-1}([a, b], \mathcal{B}([a, b])) = \mathcal{C}_b^{-1}([a, b], \mathcal{B} \cap [a, b]; \mathbb{R})$ the set of all real-valued, Borel-measurable, bounded functions on $[a, b]$. Let f be a function from $\mathcal{C}_b^{-1}([a, b], \mathcal{B}([a, b]))$ and let $P(n) = \{a_0, \dots, a_n\}$ be a partition of $[a, b]$. Introduce the following notation:

$$A_i = (a_{i-1}, a_i], \, i = 2, \dots, n, \, A_1 = [a_0, a_1],$$

$$m_i = \inf\{f(x)\colon x \in A_i\}, \, M_i = \sup\{f(x)\colon x \in A_i\},$$

$$L(f, P) = \sum_{i=1}^n m_i(a_i - a_{i-1}) \text{ (the lower Darboux sum)},$$

$$U(f, P) = \sum_{i=1}^n M_i(a_i - a_{i-1}) \text{ (the upper Darboux sum)},$$

$$l(f, P) = \sum_{i=1}^n m_i 1_{A_i}, \, u(f, P) = \sum_{i=1}^n M_i 1_{A_i}.$$

Now, a function $f : [a, b] \to \mathbb{R}$ is called a *step function*, if there is a partition P of interval $[a, b]$ such that f is constant on each subinterval (a_{i-1}, a_i), $i = 1, \dots, n$. Recall that, unlike a step function, a simple function is constant on each Borel subset (not just subinterval, of $[a, b]$) in a finite partition of $[a, b]$.

Denote by $\mathsf{S}([a, b]) = \mathsf{S}([a, b], \mathcal{B}([a, b]), \mathbb{R})$ the space of all step functions on $[a, b]$. Clearly, step functions are Borel-measurable functions and l and u are elements of $\mathsf{S}([a, b])$. Thus, $L(f, P)$ and $U(f, P)$ can be expressed in terms of Lebesgue integrals as

$$L(f, P) = \int l(f, P) d\lambda = \int\limits_{[a,b]} l(f, P) d\lambda$$

and

$$U(f, P) = \int u(f, P) d\lambda = \int\limits_{[a,b]} u(f, P) d\lambda$$

(in agreement with Notations 1.13).

Now let $\{P(n) = P(a_0, \dots, a_n);\ n = 1, 2, \dots\}$ be a sequence of partitions of $[a, b]$ such that $\{P(n), \preceq\}$ is a chain. Denote by $|P(n)|$ the Lebesgue measure of the largest subinterval of $P(n)$ and call it the *mesh* of this partition. A chain $\{P(n), \preceq\}$ is said to be *canonic*, if $\{|P(n)|\}$ is a monotone nonincreasing sequence, vanishing when $n \to \infty$.

Let $l_n = l(f, P(n))$ and $u_n = u(f, P(n))$ denote the lower and the upper step functions corresponding to partition $P(n)$ in a canonic chain. Then it can be easily verified that

$$u_1 \geq u_2 \geq \cdots \geq f \geq \dots\ l_2 \geq l_1\,. \tag{3.0}$$

Let $U_n = U(f, P(n)) = \int u_n\, d\lambda$ and $L_n = L(f, P(n)) = \int l_n d\lambda$. By monotonicity of the Lebesgue integral, we have

$$U_1 \geq U_2 \geq \cdots \geq \cdots \geq\ L_2 \geq L_1.$$

Because f is bounded, there exist $U\_ = \inf U_n = \lim U_n$ (called the *upper Darboux integral*) and $L_+ = \sup\ L_n = \lim\ L_n$ (called the *lower Darboux integral*).

3.1 Definition. If $U_- = L_+$, then their common value, $R(f,[a,b])$, is called the *Riemann integral* of the (bounded) function f over $[a,b]$, and the function f is called *Riemann integrable.* $R(f,[a,b])$ is also denoted by the symbol

$$(\mathrm{R})\int_{[a,b]} f(x)dx.$$

Sometimes, to tell a Lebesgue integral from a Riemann integral we write

$$(\mathrm{L})\int_{[a,b]} f(x)dx = \int_{[a,b]} f(x)\lambda(dx).$$

For notational consistency, most often we use the $d\lambda$ symbol within the Lebesgue integral (rather than an "L" in front of it). However, many textbooks and papers routinely use the same symbol dx in Lebesgue integrals as in Riemann integrals, which we do not believe should cause any serious confusion. □

3.2 Theorem. *Let* $f \in \mathcal{C}_b^{-1}([a,b],\mathcal{B}([a,b]))$. *If* f *is Riemann integrable on* $[a,b]$, then $f \in L^1([a,b],\mathcal{B}([a,b]),\lambda;\mathbb{R})$, *that is, it is Lebesgue integrable on* $[a,b]$. *In this case, the Riemann integral of* f *equals the Lebesgue integral of* f.

Proof. Let f be Riemann integrable. Then, $\lim_{n\to\infty}(U_n - L_n) = 0$. Applying Fatou's lemma 2.4 to the sequence $\{u_n - l_n\} \subseteq \mathcal{C}_+^{-1}([a,b],\mathcal{B}([a,b]))$, we have

$$0 \le \int \underline{\lim}\,(u_n - l_n)d\lambda = \int \lim\,(u_n - l_n)\,d\lambda$$

$$\le \underline{\lim} \int (u_n - l_n)\,d\lambda = \lim(U_n - L_n) = 0. \tag{3.2}$$

Because of (3.0), (3.2), and the fact that $u_n - f$ and $f - l_n$ are elements of $\mathcal{C}_+^{-1}$, by Lemma 1.15 we have that

$$u = \lim u_n = l = \lim l_n = f \text{ a.e.} \tag{3.2a}$$

Also, since $f \in \mathcal{C}_b^{-1}([a,b],\mathcal{B}([a,b]))$ and $l_n \le f$, it follows that

$$l_n \in L^1([a,b], \mathcal{B}([a,b]), \lambda; \mathbb{R}).$$

Now we can apply Lebesgue's dominated convergence theorem 2.6 to the sequence $\{l_n\}$ with respect to its $a.e.$-limit-function f to have

$$\lim_{n\to\infty} \int_{[a,b]} l_n \, d\lambda = \lim_{n\to\infty} L_n = (\mathrm{R}) \int_{[a,b]} f(x)dx = \int_{[a,b]} f d\lambda.$$

□

3.3 Remarks.

(i) The functions $u = \inf u_n$ and $l = \sup l_n$ in (3.0) are called the *upper* and the *lower Baire functions*. Therefore, L_+ is the Lebesgue integral of the lower Baire function l.

(ii) The above construction of the Riemann integral, which is now common in mathematical analysis courses, belongs to Gaston Darboux in his work *Mémoire sur la théorie des fonctions discontinues* of 1875. The original construction of the Riemann integral, that goes back to Augustin-Louis Cauchy (1789 – 1857) in 1823 (and later generalized by Riemann in his *Habilitationschrift* of 1854), is as follows. Given a function $f \in \mathcal{C}_b^{-1}([a,b], \mathcal{B}([a,b]))$ and a partition P from a canonic sequence of partitions $\{P(n) = P(a_0, \dots, a_n); n = 1, 2, \dots\}$ of the interval $[a,b]$, define the *Cauchy sum* as

$$C(f,P) = \textstyle\sum_{i=1}^n f(\xi_i)(a_i - a_{i-1}), \tag{3.3}$$

where ξ_i (called a *tag*) is any point of $[a_{i-1}, a_i]$. Note that, unlike the Darboux sum, the Cauchy sum is not specified because ξ_i's are arbitrary. If the limit

$$\mathrm{C} = \lim_{|P(n)|\to 0} C(f, P(n)) \tag{3.3a}$$

exists as a unique number, then f is called *Cauchy integrable on* $[a,b]$ and the value of this limit is denoted by $(\mathrm{C})\int_a^b f(x)dx$. Clearly,

$$L_+ \le \mathrm{C} \le U_-.$$

and therefore the Cauchy integral exists if $L_+ = U_-$. In 1875, Darboux proved that this is also a necessary condition for the existence of C and in this case, C = R. Darboux's theorem and his approach are subjects to standard texts in mathematical analysis, whereas Riemann's concept of the integral is more common in calculus classes as it leads to a quicker and more lucid interpretation. As a sufficient condition of the existence of the Riemann integral, Cauchy required that f be continuous on $[a, b]$. Riemann relaxed Cauchy's integrability condition by requiring that for each $\varepsilon > 0$, there is a partition P of $[a, b]$ such that $U_-(f, P) - L_+(f, P) < \varepsilon$. However, Riemann did not specify the class of functions, which are integrable (although he pointed out that a function can be discontinuous on a dense set and nevertheless integrable). Lebesgue did this in his widely referred to Theorem 3.5 which is to follow. □

3.4 Example. Let f be the Dirichlet jump function introduced in Example 1.4. Consider its modification

$$f(x) = \mathbf{1}_{\mathbb{Q}\cap[0,1]}(x) \in \mathcal{C}_+^{-1}([0,1], \mathcal{B}([0,1])).$$

The Lebesgue integral of f exists and equals zero. The Riemann integral of f, however, does not exist, in as much as for every partition, the lower Baire function equals 0 ($l = 0$) and the upper Baire function equals 1 ($u = 1$). Therefore, the lower Darboux integral $L_+ = 0$, and the upper Darboux integral $U_- = 1$. □

3.5 Theorem (H. Lebesgue). *Let* $f \in \mathcal{C}_b^{-1}([a, b], \mathcal{B}([a, b]))$. *Then* f *is Riemann integrable on* $[a, b]$ *if and only if* f *is continuous* λ-a.e. *on* $[a, b]$.

Proof.

(i) If f is continuous on $[a, b]$, then it is uniformly continuous on $[a, b]$. This implies that for each $\varepsilon > 0$, there is a $\delta > 0$, such that for each partition P whose mesh is less than δ,

$$u(f, P) - l(f, P) < \varepsilon\,. \tag{3.5}$$

(Show it; see Problem 3.1.) This leads to Riemann integrability.

(ii) Let f be bounded, Borel-measurable, and λ-a.e. continuous on $[a, b]$. If f is not continuous everywhere, but is bounded, it can have only discontinuities of finite magnitude. From the nature of the lower and the upper Baire functions, l and u, it follows that l and u coincide with f at all points of continuity of f. (A rigorous proof of this statement, known as Baire's theorem, is contained in many standard analysis textbooks.) At the points of discontinuity of f, l assumes the smallest values and u takes the largest values (this can be shown by elementary methods; see Figure 3.1.)

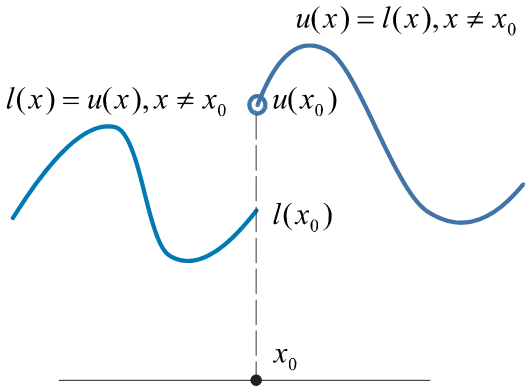


Figure 3.1

Then, if f is discontinuous on a negligible set S, it should follow that u and l differ on the same set S. By the above condition, $S \subseteq N$ where N is a measurable null set. Because f is bounded, u_n and l_n are measurable, bounded jump functions, and U_n and L_n exist. By the Lebesgue dominated convergence theorem, $U_{\_} - L_{+} = 0$, which implies that f is Riemann integrable. Indeed, by Lemma 1.15 (because $l_n \leq u_n$),

$$U_{-} - L_{+} = \lim_{n \to \infty} U_n - L_n = \lim_{n \to \infty} \int (u_n - l_n)\, d\lambda$$

$$\int (u - l) d\lambda = 0,$$

because $u - l = 0$ on N^c, i.e., $u = l$ λ-$a.e.$

(iii) Let f be Riemann integrable. Then, by Theorem 3.2 equation (3.2a) holds:

$$f = l = \lim_{n\to\infty} l_n \ = u = \lim_{n\to\infty} u_n \quad \text{a.e.}$$

Furthermore, f is bounded. We utilize again the above arguments. From the nature of u and l, it follows that, in this case, $u, l,$ and f coincide wherever f is continuous. At all points of discontinuity, whereas f assumes one of these values, the smallest values of them are assigned to l and the largest ones to u. Therefore, the set, on which the function f is discontinuous, equals the set on which u and l differ. This proves that f is continuous λ-a.e. □

3.6 Remarks.

(i) By employing a canonic chain of partitions on the X-axis, in construction of the Riemann integral we sometimes face the problem that the sequence of the corresponding lower jump functions $\{l_n\}$ converges to the lower Baire function l, but it does not converge to f, as is the case for the Dirichlet function. Consequently, the lower Darboux integral gives a "wrong" value. In contrast, the construction of the Lebesgue integral literally sets up partitions on the Y-axis whose canonic chains form monotone increasing sequences of "lower" jump functions. The latter, due to Theorem 5.5, Chapter 5, always converge to f. Consequently, the lower Darboux integral L_+ equals the Lebesgue integral $\int f d\mu$.

(ii) Although Riemann and Darboux enlarged the previously existing class of integrable functions, the Riemann integral has many limitations, one of which goes back to the fundamental theorem of calculus in the form

$$(\mathrm{R})\int_a^b f'(x)dx = f(b) - f(a).$$

This formula becomes meaningless when a differentiable function f is not integrable. On the other hand, the classical proof of the formula

$$\frac{d}{dx}\int_a^x f(u)du = f(x)$$

was originally based on the continuity assumption for f. The new concept of integration suggested by Henri Lebesgue in 1902 in his doctoral work restored the generality of the fundamental theorem to its current status. Furthermore, the class of Lebesgue integrable functions is significantly enlarged.

Notice that from Theorem 6.5, Chapter 5, it follows that, unlike the Cauchy - Riemann - Darboux formation of partitions of $[a, b]$ essentially leading to Definition 3.1 or (3.3a), in the Lebesgue integral of f, partitions on $[a, b]$ are inverse images under f^* of partitions of interval $[0, \sup f]$ on the Y-axis. This approach (based on an explicit approximation of f by simple functions) provides a better detection and incorporation of fine variations of f (which Riemann integral does not) thereby allowing the Lebesgue integral to handle a much wider class of functions on $[a, b]$. Besides, defining the Lebesgue integral through the integral for simple functions enables far more powerful and easy-to-accomplish operations manifested by convergence theorems under more relaxed conditions compared to the Riemann integral.

On the other hand, this deficiency of the Riemann integral was later on corrected by the British mathematician Ralph Henstock (1923 – 2007) in 1955 and, independently, the Czech Jaroslav Kurzweil (born 1926) in 1957 by refining ideas of the French Arnaud Denjoy in 1912 and German Oscar Perron in 1914, and leading to the *gauge* or *Henstock-Kurzweil integral* (cf. [Bart], [Gor], [Hen], [LeeP], [LeeT], and [Swa]).

In short, the gauge integral G of f on $[a, b]$ is defined as follows. Given an $\varepsilon > 0$, if there is a function $[[a, b], (0, \infty), \delta]$ (called a *gauge function*) such that, for any partition $a = a_0 \le a_1 < \cdots < a_n = b$ and any choice of tags $\xi_i \in [a_{i-1}, a_i]$ satisfying $a_{i-1} - \delta(a_{i-1}) < a_{i-1} \le \xi_i \le a_i < a_i + \delta(a_i)$, $i = 1, \ldots, n$,

$$|\mathrm{G} - \textstyle\sum_{i=1}^n f(\xi_i)(a_i - a_{i-1})| < \varepsilon$$

holds, the gauge integral G is said to exist. In the definition of the Riemann integral, the gauge function δ is constant, which is the core defect that makes the Riemann integration less flexible. There are more gauge-integrable functions on $[a, b]$ than Lebesgue- and "improper" Riemann-integrable functions combined. Some advocated even that both Riemann and Lebesgue integrals be replaced with the gauge integral in calculus and analysis courses. However, it did not bring in a broad support for many good reasons (which are beyond the scope of this book). Most importantly, the Lebesgue integral in the advanced literature on real analysis and measure theory, is often used as a mere illustration to the mighty abstract integral that dominates stochastic and functional analyses.

(iii) As we mentioned, in order that a function be Lebesgue integrable, it need not be bounded. A class of Riemann-integrable functions, as known, can be "extended" for unbounded functions by the use of the "improper integral." Another need for the improper integral arises when the interval of integration is unbounded. In the latter case, the integral is constructed as usual on a compact interval $[a, b]$, and then its values are taken for $a \to -\infty$ or $b \to \infty$. This is a trick rather than a conventional integral construction. (See the further implication of this method below.) That is why such integrals are called improper.

(iv) Unlike improper integration over infinite intervals, there is another way to integrate functions with the conventional approach of constructing an integral via uniform "partitions" of an infinite interval. Consider for example, a bounded Borel measurable function f on interval $[a, \infty)$, and partition this interval by the sequence $\{a_n\}$, where $a_n = a + \delta n,\ \ n = 0,\ 1,\ \ldots,$ for some positive δ. Then on each of the intervals $\Delta_n = [a_n, a_{n+1})$ consider

$$m_n = \inf\{f(x) : x \in \Delta_n\} \quad \text{and} \quad M_n = \sup\{f(x) : x \in \Delta_n\}.$$

Because the Lebesgue measure of each interval Δ_n equals δ, we have again the lower Darboux sum,

$$L(f,\, \delta) = \ \delta \textstyle\sum_{n\geq 0} m_n\,,$$

and the upper Darboux sum,

$$U(f,\delta) = \delta \sum_{n\geq 0} M_n \,.$$

If $\lim\limits_{\delta \downarrow 0} L(f,\delta) = \lim\limits_{\delta \downarrow 0} U(f,\delta)$ then its common value is denoted by

$$(\mathrm{D}) \int_a^\infty f(x)\,\mathrm{d}x$$

and called the *direct Riemann integral.* The function f is then said to be *directly Riemann integrable*. The direct integrability is used in probability, specifically in renewal theory, where such a notion is introduced for a class of nonnegative functions bounded over finite intervals. We not that the class of directly integrable functions is essentially more narrow than the class of Riemann integrable functions in the improper sense. □

3.7 Examples.

(i) Let $\Omega = [0,1]$ and let $f(x) = x^2 \mathbf{1}_A(x) + \sin x\, \mathbf{1}_{A^c}(x)$, where A^c is the Cantor ternary set. The function f is a bounded Borel-measurable function on $[0,1]$ and obviously λ-a.e. continuous on $[0,1]$. Thus, f is Lebesgue as well as Riemann integrable and $f(x) = x^2$ λ-a.e. on $[0,1]$. Furthermore,

$$(\mathrm{R}) \int\limits_{[0,1]} f(x)dx = (\mathrm{L}) \int\limits_{[0,1]} f(x)\lambda(dx)$$

$$= (\mathrm{L}) \int\limits_{[0,1]} x^2 \lambda(dx) = (\mathrm{R}) \int\limits_{[0,1]} x^2 dx = \tfrac{1}{3}.$$

(ii) Let $\Omega = [1,2]$ and $f(x) = (x-1)^{-1/3}$. We wish to evaluate $\int\limits_{[1,2]} f(x)\lambda(dx)$. Because f is no longer bounded (on $[1,2]$) we cannot apply the same techniques as discussed above. Consequently, we introduce an auxiliary sequence of functions, $\{f_n\}$, defined as

$$f_n(x) = \begin{cases} n, & 1 \leq x < 1 + \frac{1}{n^3} \\ (x-1)^{-1/3}, & 1 + \frac{1}{n^3} \leq x \leq 2 \end{cases}$$

(see Figure 3.2). It is easily seen that $\{f_n\}$ is a monotone increasing sequence of continuous functions contained in $\mathcal{C}_+^{-1}([1,2], \mathcal{B}([1,2]))$ with $\sup\{f_n\} = f$.

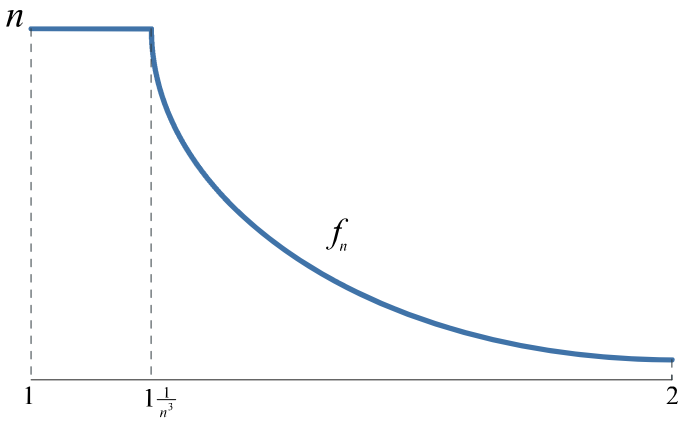


Figure 3.2

By Proposition 5.6 (iv), Chapter 5, $f \in \mathcal{C}_+^{-1}$. By the monotone convergence theorem,

$$\int_{[1,2]} f(x)\lambda(dx) = \sup \int_{[1,2]} f_n(x)\lambda(dx).$$

On the other hand,

$$\int_{[1,2]} f_n d\lambda = (\mathrm{R}) \int_{[1,2]} f_n(x)dx$$

$$= (\mathrm{R}) \int_{[1,1+1/n^3]} n\,dx + (\mathrm{R}) \int_{[1+1/n^3,2]} (x-1)^{-1/3} dx = \frac{3}{2} - \frac{1}{2} \cdot \frac{1}{n^2}.$$

Thus,

$$\sup \int_{[1,2]} f_n(x)\lambda(dx) = \frac{3}{2}.$$

Observe that the Riemann improper integration technique for unbounded functions could also be applied to this function. □

3.8 Remark. As mentioned, the Lebesgue integrable functions form a much wider class compared to the Riemann integrable functions. Also recall that an $L^1(\mathbb{R}, \mathcal{B}, \lambda)$-function f can be integrated over arbitrary Borel sets, whereas the Riemann integral is defined just on intervals. With all these benefits, however, the Lebesgue integral does not have the same elegance and analytical tractability the Riemann integral has through to its "Newton - Leibnitz bridge" to derivatives and antiderivatives and a huge reserve of integration techniques. In many cases, whenever possible, the Lebesgue integral is just reduced to a Riemann integral. In addition, the class of Riemann integrable functions is enlarged through the functions that are Riemann integrable in an improper sense. This class includes functions with discontinuities of an infinite magnitude and functions defined on intervals of type $[a, \infty)$ or $(-\infty, b]$ or $(-\infty, \infty)$.

In Example $3.7(ii)$ we examined a Lebesgue integral of an unbounded function. In a certain sense, the approach used there is reminiscent of Riemann integration of unbounded functions. In the proposition below we state that in most cases, when the integration over an infinite interval is needed, we can use Riemann integration in the improper sense and equate their values to those for Lebesgue integrals. This fact makes the Riemann improper integral fully legitimate. □

3.9 Proposition. *Let $f \in \mathcal{C}_+^{-1}(\mathbb{R}, \mathcal{B}; \mathbb{R}_+)$ let f be Riemann integrable on any compact interval. Then $f \in L^1(\mathbb{R}, \mathcal{B}, \lambda; \mathbb{R}_+)$ if and only if the improper Riemann integral of f,*

$$\mathrm{R} = \lim_{\substack{a \to -\infty \\ b \to \infty}} \int\limits_{[a,b]} f(x)dx,$$

exists. [*We say that $f \in \mathcal{R}(\mathbb{R})$, where $\mathcal{R}(\mathbb{R})$ is the class of all functions on $\mathbb{R}$ Riemann integrable in the improper sense.*] *In this case* $\mathrm{R} = \int f d\lambda$.

Proof. Denote

$$R_{nk} = (\mathrm{R}) \int\limits_{B_{nk}} f(x)dx \text{ where } B_{nk} = [-k, n].$$

Then, because f is Riemann integrable, $R_{nk} = \int f\mathbf{1}_{B_{nk}} d\lambda$. Observing that

$$f = \sup\{f\mathbf{1}_{B_{nk}} : n = 1, 2, \dots;\ k = 1, 2, \dots\},$$

we have, by the monotone convergence theorem,

$$\int f\, d\lambda = \sup R_{nk} = \mathrm{R} < \infty. \qquad \square$$

3.10 Remark. The special case treated in the above proposition and applied to nonnegative Borel measurable functions can easily be extended to arbitrary functions of $\mathcal{C}^{-1}$ by noticing that $f \in L^1(\mathbb{R}, \mathcal{B}, \lambda; \overline{\mathbb{R}})$ if and only if $|f| \in L^1(\mathbb{R}, \mathcal{B}, \lambda; \overline{\mathbb{R}})$. Therefore, using Proposition 3.9, we conclude that $|f|$ must be an element of $\mathcal{R}(\mathbb{R})$. In this case, evidently,

$$\begin{aligned} &(\mathrm{R}) \textstyle\int_{-\infty}^{\infty} f(x)\, dx \\ &= (\mathrm{R}) \textstyle\int_{-\infty}^{\infty} f^+(x)\, dx - (\mathrm{R}) \int_{-\infty}^{\infty} f^-(x)\, dx \\ &= \textstyle\int f^+\, d\lambda - \int f^-\, d\lambda = \int f\, d\lambda. \end{aligned} \qquad \square$$

3.11 Examples.

(i) Consider the function $f(x) = \frac{x \sin x}{k^2 + x^2}$ (where $k \neq 0$). We show that this function is Riemann integrable in the improper sense but not Lebesgue integrable over $\mathbb{R}_+$. We apply the Dirichlet criterion:

> *Let g and h be two real-valued functions defined on $[a, \infty)$. If g is monotonically vanishing at ∞ and $\left|\, (\mathrm{R}) \int_a^b h(x)dx \,\right| \leq C$, for each $b > a$ and positive real number C (i.e., the integral of h is uniformly bounded in b), then the improper integral $(\mathrm{R}) \int_a^{\infty} gh$ is convergent.*

In our case, the function $\frac{x}{k^2+x^2}$ can be taken for g and $\sin x$ can represent h. Then, the conditions in Dirichlet's criterion are met for $a = 0$, and consequently, $(\mathrm{R})\int_0^\infty f$ converges. On the other hand, $f \in L^1(\mathbb{R}, \mathcal{B}, \lambda; \mathbb{R})$ if and only if $|f| \in L^1(\mathbb{R}, \mathcal{B}, \lambda; \mathbb{R})$, which, by Proposition 3.9, is equivalent to the convergence of the integral $(\mathrm{R})\int_0^\infty |f| := I$. It is shown that $I = \infty$. Indeed,

$$I = \sum_{n=0}^\infty \int_{n\pi}^{(n+1)\pi} \frac{|\sin x| x}{k^2+x^2}\, dx$$

$$= \sum_{n=0}^\infty \int_0^\pi \frac{\sin t(t+n\pi)}{k^2+(n\pi+t)^2}\, dt \geq \frac{\pi n \int_0^\pi \sin t\, dt}{k^2+\pi^2(n+1)^2}$$

(the second summation is due to the inequality $\pi n + t \leq (n+1)\pi$, for $t \in [0, \pi]$).

Thus

$$I \geq \sum_{n=0}^\infty \frac{2\pi n}{k^2+\pi^2(n+1)^2} = \infty.$$

(ii) The function $f(x) = \sin x \exp(-\frac{x^2}{2})$ is an element of $\mathcal{C}^{-1}$ and it is Lebesgue integrable, because $|f(x)| \leq g(x) = \exp(-\frac{x^2}{2})$ and $g(x) \geq 0$ and because

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x)dx = 1\ .$$

Observe that $x \mapsto \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$, $x \in \mathbb{R}$, is the normal density function of the standard normal distribution. [See Example 5.10 (iii).] □

PROBLEMS

3.1 Prove (3.5) in Theorem 3.5.

3.2 In Example 3.4, we showed that the Dirichlet function f on $[0, 1]$ is Lebesgue integrable, but not Riemann integrable. Because the rationals have the Lebesgue measure of 0, the

function f is equal to 0 (a constant) for λ-almost all points on $[0,1]$, and therefore, it is continuous almost everywhere on $[0,1]$. By Theorem 3.5, f must be Riemann integrable. This is just the opposite of the result of Example 3.4. What is wrong with this reasoning?

3.3 Is the function $f(x) = \frac{1}{x}$ on $[0,1]$ Borel-measurable and λ-integrable? Is $f \in \mathbb{L}([0,1], \mathcal{B}([0,1], \lambda))$?

3.4 Show that the function f, such that $f(x) = \frac{1}{x}\cos(\frac{1}{x})$ on $(0,1]$ and $f(0) = 0$, is Borel-measurable and not λ-integrable on $[0,1]$.

3.5 Let $f\colon [0,1] \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} (-1)^{n+1}n, & x \in A_n = (\frac{1}{n+1}, \frac{1}{n}],\ n = 1,2,\dots \\ 0, & x = 0. \end{cases}$$

Show that f is improperly Riemann integrable but not Lebesgue integrable.

3.6 Let f be a monotone increasing differentiable function on $[a,b]$ and let φ be its inverse function on $[f(a), f(b)]$. Prove that

$$\int_a^b f(x)\lambda(dx) = \int_{f(a)}^{f(b)} y\varphi'(y)\lambda(dy).$$

3.7 Investigate if the function $f(x) = \frac{\sin x}{x^a}\,\mathbf{1}_{\{x\neq 0\}}(x)$ (where $0 < a \leq 1$) is improperly Riemann and Lebesgue integrable over $\mathbb{R}_+$.

3.8 Let G be a nonempty open subset of $[a,b]$ and let f be a Borel measurable function on $[a,b]$, discontinuous at each point of G. Can f be Riemann integrable?

3.9 Show that the functional

$$\|f\|_{L^1} = \int_a^b |f| d\lambda$$

defines a seminorm on $L^1([a,b], \mathcal{B}([a,b]), \lambda)$. How can $\|\cdot\|_{L^1}$ be "made" a norm?

3.10 Let $s \in \Psi_+([a,b], \mathcal{B}([a,b]))$. Show that for each $\varepsilon > 0$, there is a continuous function $h \in \mathcal{C}([a,b])$ such that

$$\| s - h \|_{L^1} < \varepsilon.$$

3.11 Show that the space $\mathcal{C}([a,b])$ of all continuous functions on interval $[a,b]$ is dense in $(L^1([a,b], \mathcal{B}([a,b]), \lambda), \|\cdot\|_{L^1})$.

3.12 Use Lebesgue's Theorem 3.5 to show that the limit of a uniformly convergent sequence $\{f_n\}$ of bounded Riemann integrable functions on $[a,b]$ is Riemann integrable on $[a,b]$. Prove that under this condition,

$$\lim_{n\to\infty} (\mathrm{R}) \int_a^b f_n(x) dx = (\mathrm{R}) \int_a^b \lim_{n\to\infty} f_n(x) dx. \quad \text{(P3.12)}$$

3.13 Let A be a subset of $[a,b]$ whose closure is negligible w.r.t. Borel - Lebesgue measure λ. Is the function $g = \mathbf{1}_A$ Riemann integrable on $[a,b]$? If it is Riemann integrable, then calculate $(\mathrm{R}) \int_a^b g(x) dx$.

3.14 Let $\{f_n\}$ be a sequence of bounded, Borel measurable, nonnegative functions on $A \subseteq \mathbb{R}$. Suppose $(\mathrm{L}) \int_A f_n d\lambda \to 0$ for $n \to \infty$. Is it true that $f_n \to 0$ λ-a.e. on A?

NEW TERMS:

Riemann, Bernhard, 399
Cauchy, Augustin-Louis, 399
Darboux, Jean-Gaston, 399
partition, 399
refinement, 399
$\mathcal{C}_b^{-1}([a,b],\mathcal{B}([a,b]);\mathbb{R})$, Borel-measurable bounded functions, 399
$L(f,P)$, lower Darboux sum, 399
$U(f,P)$, upper Darboux sum, 399
step function, 400
$\mathsf{S}([a,b],\mathcal{B}([a,b]),\mathbb{R})$, space of step functions, 400
mesh of a partition, 400
canonic chain of partitions, 400
upper Darboux integral, 400
lower Darboux integral, 400
Riemann integral, 401
Riemann integrable function, 401
criterion of Lebesgue integrability, 401
upper Baire function, 402
lower Baire functions, 402
Cauchy sum, 402
Cauchy integrable function, 402
Dirichlet function, 403
Lebesgue's Theorem of Riemann integrability, 403
Henstock, Ralph, 406
Kurzweil, Jaroslav, 406
Denjoy, Arnaud, 406
Perron, Oscar, 406
gauge integral, 406
Henstock-Kurzweil integral, 406
gauge function, 406
improper Riemann integral, 407, 409
direct Riemann integral, 408
direct Riemann integrability, 408
Dirichlet's criterion, 411
$(L^1([a,b],\mathcal{B}([a,b]),\lambda),\|\cdot\|_{L^1})$, L^1-seminormed space, 414

4. INTEGRATION WITH RESPECT TO IMAGE MEASURES

In this short section we challenge the claim (on a small scale though) that the Riemann integral has an overwhelming variety of techniques making it superior to the Lebesgue integral. Besides, we focus on the abstract integral. As one of the extensions of major integration techniques, we study integration with respect to image measure μF^* (where F is a measurable mapping), under the nickname *change of variables*, as it resembles the prominent method for the Riemann integral. A more specific approach to a change of variables for Lebesgue integrals in n-dimensional Euclidean spaces is treated separately in [Dsh2].

4.1 Theorem (Change of Variables). *Let $(\Omega_0, \Sigma_0, \mu)$ be a measure space, $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$, and $F\colon (\Omega_0, \Sigma_0) \to (\Omega, \Sigma)$ be a measurable map [such that μF^* is an image measure on the measurable space (Ω, Σ)]. Then, the following formula holds true.*

$$\int f(\omega)\, d\mu F^*(\omega) = \int f \circ F(\omega_0) d\mu(\omega_0). \tag{4.1}$$

Specifically, if $f = g\mathbf{1}_A$, where $A \in \Sigma$ and $g \in \mathcal{C}^{-1}(\Omega, \Sigma)$, then (4.1) *reduces to*

$$\int\limits_A g(\omega) d\mu F^*(\omega) = \int\limits_{F^*(A)} g \circ F(\omega_0)\, d\mu(\omega_0). \tag{4.1a}$$

Proof.

(i) Let $s \in \Psi_+(\Omega, \Sigma)$ be just an indicator function $s = \mathbf{1}_A$. By Problem 3.7, Chapter 1, we have that

$$s \circ F(\omega_0) = \mathbf{1}_A \circ F(\omega_0) = \mathbf{1}_{F^*(A)}(\omega_0).$$

Therefore,

$$\int \mathbf{1}_A \circ F(\omega_0)\, d\mu(\omega_0) = \int \mathbf{1}_{F^*(A)}(\omega_0)\, d\mu(\omega_0)$$

$$= \mu(F^*(A)) = \mu F^*(A) = \int \mathbf{1}_A(\omega) d\mu F^*(\omega).$$

(ii) Let s be a nonnegative simple function with the representation,

$$s = \sum_{i=1}^{n} a_i \, \mathbf{1}_{A_i} \, , \, a_i \in \mathbb{R}_+.$$

Then,

$$\begin{aligned} s \circ F(\omega_0) &= \Big(\sum_{i=1}^{n} a_i \, \mathbf{1}_{A_i} \Big) \circ F(\omega_0) \\ &= \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} \circ F(\omega_0) = \sum_{i=1}^{n} a_i \mathbf{1}_{F^*(A_i)}(\omega_0) \end{aligned}$$

and

$$\int s \circ F \, d\mu = \sum_{i=1}^{n} a_i \, \mu F^*(A_i) = \int s d\mu F^*.$$

(iii) Let $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$. Then there exists $\{s_n\} \uparrow \subseteq \Psi_+$ such that $f = \sup\{s_n\}$. For s_n we have, according to (ii):

$$\int s_n \circ F \, d\mu = \int s_n \, d\mu F^*.$$

Observe that $\{s_n \circ F\} \uparrow \subseteq \Psi_+(\Omega_0, \Sigma_0)$ and, by Proposition 5.6(iv),

$$\sup\{s_n \circ F\} = f \circ F \in \mathcal{C}_+^{-1}(\Omega_0, \Sigma_0).$$

Therefore, we have that

$$\begin{aligned} \int f \circ F \, d\mu &= \sup \{ \int s_n \circ F \, d\mu \} \\ &= \sup\{\int s_n \, d\mu F^*) = \int f \, d\mu F^*. \end{aligned}$$

(iv) Let $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$. Then, $f = f^+ - f^-$ and, according to Problem 4.1,

$$(f \circ F)^+ = f^+ \circ F \text{ and } (f \circ F)^- = f^- \circ F.$$

Therefore,

$$f \circ F = (f \circ F)^+ - (f \circ F)^- = f^+ \circ F - f^- \circ F,$$

and this, along with (iii), implies that

$$\begin{aligned} \int f \circ F \, d\mu &= \int (f \circ F)^+ \, d\mu - \int (f \circ F)^- \, d\mu \\ &= \int f^+ \circ F \, d\mu - \int f^- \circ F \, d\mu \\ &= \int f^+ \, d\mu F^* - \int f^- \circ F^* \, d\mu = \int f \, d\mu F^*. \end{aligned}$$

(v) Let $f = g\mathbf{1}_A$ where $A \in \Sigma$ and $g \in \mathcal{C}^{-1}(\Omega, \Sigma)$. Then we have

$$\begin{aligned} \int_A g \, d\mu F^* &= \int g\mathbf{1}_A \, d\mu F^* = \int (g \circ F)(\mathbf{1}_A \circ F) \, d\mu \\ &= \int g \circ F \, \mathbf{1}_{F^*(A)} d\mu = \int_{F^*(A)} g \circ F d\mu. \end{aligned}$$ □

4.2 Corollary. *Let (Ω, Σ, μ) be a measure space and let*

$$F\colon (\Omega, \Sigma) \to (\Omega, \Sigma)$$

be a bijective transformation that is Σ-Σ measurable along with its inverse F^*. Then, for each $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$, the following formula holds true.

$$\int_{F^*(A)} f \, d\mu = \int_A f \circ F^* \, d\mu F^*. \tag{4.2}$$

(See Problem 4.2.)

4.3 Example. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and let $X \in \mathcal{C}^{-1}(\Omega, \Sigma)$ be a r.v.. Recall that X induces the image measure $\mathbb{P}X^*$, or, equivalently, the probability distribution on the measurable space $(\overline{\mathbb{R}}, \overline{\mathcal{B}})$, thereby generating the new probability space $(\overline{\mathbb{R}}, \overline{\mathcal{B}}, \mathbb{P}X^*)$. The functional of X, $\int X \, \mathbb{P}$, was called [in Definition 1.10(iv)] the *expectation of the r.v. X* and denoted by symbol $\mathbb{E}X$.

Let $g \in \mathcal{C}^{-1}(\overline{\mathbb{R}}, \overline{\mathcal{B}})$. Then, $g \circ X$ is also a r.v. whose expectation can be expressed as

$$\mathbb{E}[g \circ X] = \int g \circ X(\omega)\, \mathbb{P}(d\omega).$$

By formula (4.1), we have

$$\mathbb{E}[g \circ X] = \int g(x)\, \mathbb{P}X^*(dx). \tag{4.3}$$

Specifically, if $g(x) = x$, we have $\mathbb{E}X = \int x \mathbb{P}X^*(dx)$. If $g = \mathbf{1}_A$, then from (4.3), Notation 1.13, and Definition 4.2, Chapter 5,

$$\mathbb{E}[\mathbf{1}_A \circ X] = \int_A \mathbb{P}X^*(dx) = \mathbb{P}X^*(A) = \mathbb{P}\{X \in A\}. \tag{4.3a}$$

□

PROBLEMS

4.1 Show that $(f \circ F)^+ = f^+ \circ F$ and $(f \circ F)^- = f^- \circ F$.

4.2 Prove Corollary 4.2: *Let* (Ω, Σ, μ) *be a measure space and let*

$$F\colon (\Omega, \Sigma) \to (\Omega, \Sigma)$$

be a bijective transformation which is Σ-Σ *measurable along with its inverse* F^*. *Then, for each* $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$, *the following formula holds true.*

$$\int_{F^*(A)} f\, d\mu = \int_A f \circ F^*\, d\mu F^*. \tag{4.2}$$

4.3 Let $f \in \mathcal{C}^{-1}(\mathbb{R}^n, \mathcal{B}^n)$, $\alpha \in \mathbb{R}$, $b \in \mathbb{R}^n$, and $B \in \mathcal{B}^n$. Show that

$$\int_B f(\alpha x + b)\lambda(dx) = \tfrac{1}{|\alpha|^n} \int_A f(x)\lambda(dx),$$

where $A = \alpha B + b$.

4.3 Simplify $\int\limits_A f(e^{2x})\lambda(dx)$, where $f \in \mathcal{C}^{-1}(\mathbb{R}, \mathcal{B}; \overline{\mathbb{R}})$ and $A = [1, 2]$.

4.4 Use the change of variables formula to evaluate the integral $\int\limits_A f(2x+1)\lambda(dx)$, where

$$f(x) = \begin{cases} x^2, & x \in \mathbb{Q}^c \\ \infty, & x \in \mathbb{Q} \end{cases}$$

and $A = [1, 3]$.

NEW TERMS:

change of variables formula, 416
change of variables for a bijective transformation, 418
expectation of a random variable, 418
expectation of a function of a random variable, 419
change of variables formula with respect to the affine transformation, 419

5. POSITIVE MEASURES GENERATED BY INTEGRALS. ABSOLUTE CONTINUITY

In this section we learn that the integral $\int_A f d\mu$, as a set function, say $\nu(A)$, turns out to be a measure. Hence, the two measures, μ (the original measure) and ν (generated by the integral), are related through the given integrand-function f, which is referred to as a *density*. Conversely, under what condition imposed on two arbitrary given measures can a density function (that binds them through the integral) exist? The question raised leads to one of the central results in measure theory and integration, known as the Radon - Nikodym theorem, which specifies exactly that condition.

This section gives a very brief and informal acquaintance with the Radon - Nikodym theorem and its ramifications needed to advance to the upcoming material and serving as an introduction. A more elaborated and general version of the Radon - Nikodym theorem is treated in Section 8.

Let (Ω, Σ, μ) be a measure space. Consider the integral

$$A \to \int_A f d\mu = \int f \mathbf{1}_A \, d\mu$$

as a set function on Σ. If $f \geq 0$, then as the following proposition asserts, we have a measure on Σ.

5.1 Proposition. *Let (Ω, Σ, μ) be a measure space and let $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$. Then, the set function $\nu(A) = \int_A f d\mu$ is a measure on Σ.* □

(See Problem 5.1.)

5.2 Definition. According to Proposition 5.1, ν is a measure generated by integral $\int f \, d\mu$; ν is also called the *indefinite integral of f with respect to μ.* The function f is called a (*Radon - Nikodym*) *density function of ν relative to μ.* □

The following examples motivate the Radon - Nikodym theorem.

5.3 Example.

(i) Let (Ω, Σ, μ) be a measure space and let μ be σ-finite. Then, there exists a sequence $\{A_n\} \uparrow \Omega$ such that $\mu(A_n) < \infty$. Define the sequence $\{a_n\} \subseteq \mathbb{R}_+ \backslash \{0\}$ as

$$a_n = \min\Big\{ \tfrac{1}{\mu(A_n)2^n} \, , \, \tfrac{1}{2^n} \Big\}, \;\; n = 1, 2, \dots \; .$$

Let

$$g_n = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} \;\; \text{and} \;\; g = \sup\{g_n\} = \sum_{i=1}^{\infty} a_i \mathbf{1}_{A_i}.$$

Then,

$$\textstyle\int g \, d\mu \; = \sup\{ \int g_n \, d\mu\} = \sum_{i=1}^{\infty} a_i \mu(A_i) \leq \sum_{i=1}^{\infty} \frac{1}{2^i} = 1.$$

Therefore, if μ is σ-finite, there always exists a positive element g from $L^1(\Omega, \Sigma, \mu)$. Conversely, let $g > 0$ and$d$ $g \in L^1(\Omega, \Sigma, \mu)$. Then

$$A_n = \{g \geq \tfrac{1}{n}\} \in \Sigma$$

and $gn \geq \mathbf{1}_{A_n}$. Thus

$$\textstyle\mu(A_n) \; = \int \mathbf{1}_{A_n} d\mu \leq n \int g \, d\mu < \infty,$$

which implies that $\mu(A_n) < \infty$. Since $g > 0,$ it follows that $A_n \uparrow \; \Omega$.

(ii) We have shown that σ-finiteness of μ is equivalent to the existence of a positive integrable function g. In other words, there is a positive "Radon - Nikodym density" g such that the measure ν generated by the integral is finite. Another noteworthy observation is that if

$$0 = \nu(A) = \textstyle\int g \mathbf{1}_A d\mu \, ,$$

then, because $g\mathbf{1}_A \geq 0$, by Lemma 1.15, $g\mathbf{1}_A \in [\mathbf{0}]_\mu$. In as much as $g > 0, \; A \in \mathcal{N}_\mu,$ i.e., from $\nu(A) = 0$ it follows that $\mu(A) = 0$.

On the other hand, should $\mu(A) = 0$, then (due to the convention $0 \cdot \infty = 0$) $g\mathbf{1}_A \in [\mathbf{0}]_\mu$ and $\nu(A) = 0$. Thus, $\nu(A) = 0$ if and only if $\mu(A) = 0$. In other words, ν and μ possess the same null sets.

However, if g is just nonnegative, $\nu(A) = 0$ does not necessarily imply that $\mu(A) = 0$. But from $\mu(A) = 0$, it follows anyway that $\nu(A) = 0$. □

We saw two different ways when measures and densities are related. Firstly, given a nonnegative measurable function f and a measure μ we have measure $\nu = \int f d\mu$. Secondly, in the above examples, given a σ-finite measure μ, we designed an integrable function $g > 0$ such that $\nu = \int g d\mu$ was a measure. In both cases, a μ-null set is also a ν-null set, in notation $\nu \ll \mu$.

In general, if ν and μ are two measures such that $\nu \ll \mu$ does a density function f of ν relative to μ always exist? The answer is given in the Radon - Nikodym theorem below.

5.4 Definition. Let μ and ν be two measures on a measure space (Ω, Σ). The measure ν is called (*absolutely*) *continuous with respect to* μ if every μ-null set is also a ν-null set. In this case, we write $\nu \ll \mu$. Any Borel measure continuous with respect to the Lebesgue measure is just called *continuous*. □

The use of the word "continuity" is due to the following proposition.

5.5 Proposition. *Let ν be a finite measure on (Ω, Σ) and let μ be another measure on (Ω, Σ). Then the following are equivalent*:

(A) $\nu \ll \mu$.

(B) *For all $\varepsilon > 0$, there is $\delta > 0$, such that for each $A \in \Sigma$ with $\mu(A) < \delta$, the inequality $\nu(A) < \varepsilon$ holds*.

Proof.

(i) Suppose statement (B) is true. Choose an ε. Denote by Δ the set of all $A \in \Sigma$, for which $\mu(A) < \delta$. Then $\mathcal{N}_\mu \subseteq \Delta$ (where $\mathcal{N}_\mu$ denotes the subset of all μ-null sets). Then, for all $N \in \mathcal{N}_\mu$, $0 = \mu(N) < \delta$ and $\nu(N) < \varepsilon$. Because ε can be made arbitrarily small, we conclude that $\nu(N) = 0$ and thus $\nu \ll \mu$.

(ii) Suppose now that statement (B) is not true. That means, for some $\varepsilon > 0$ and for any $\delta > 0$ there is a set $A(\delta) \in \Sigma$ such that $\mu(A(\delta)) < \delta$ implies that $\nu(A(\delta)) > \varepsilon$. We now define the sequence of δ's as $\delta_n = \frac{1}{2^n}$, $n = 1, 2, \dots,$ and construct the corresponding sequence of A's such that $A(\delta_n) = A_n$ with the above property, that is, $\{A_n\}$ is a μ-monotone decreasing sequence but "ν-resistant." Let $A = \overline{\lim}\, A_n$. Then $A \subseteq \bigcup_{m=n}^{\infty} A_m$ and

$$\mu(A) \le \mu\Big(\bigcup_{m=n}^{\infty} A_m\Big) \le \sum_{m=n}^{\infty} \mu(A_m) < \tfrac{1}{2^{n-1}},\ n = 1, 2, \dots\ .$$

Therefore, $\mu(A) = 0$. However, by Problem 2.5, because ν is finite,

$$\nu(A) = \nu(\overline{\lim}\, A_n) \ge \overline{\lim}\, \nu(A_n) \ge \varepsilon > 0$$

and thus ν is not μ-continuous. Hence (A) is not true either. □

The most general version of the celebrated Radon - Nikodym theorem was proved by the Pole Otto Nikodym in his paper, *Sur une généralisation des intégrales de M.J. Radon* of 1930. Another prominent Pole, Stanislav Saks, suggested the name of this theorem, perhaps meaning Nikodym's theorem on Radon integrals, although Radon himself proved a much more special case.

The idea of Radon - Nikodym's result had its inception in an 1884 paper by Thomas Stieltjes, in which he introduced the new concept of a density function in connection with his famous "Stieltjes integral" (in its present version known as the Riemann - Stieltjes integral) and initially applied to very restricted classes of functions. In 1909, Frigyes Riesz proved in his widely referred to representation theorem that Stieltjes integrals are represented by the most general continuous linear functionals on $[a, b]$ (whose more general version is explored in [Dsh2]).

Riesz's result yielded many generalizations, of which the most fruitful was by Johann Radon [Rad] in his 1913 paper, "Theorie und Anwendungen der absolut additiven Mengenfunktionen." In this paper, Radon, combining the ideas of Lebesgue and Riesz, intro-

duced an integral with respect to Borel measures on the Borel σ-algebra in $\mathbb{R}^n$ rather than the Borel - Lebesgue measure used by Lebesgue. Among other things, Radon showed the existence of a Radon - Nikodym density function with respect to this integral as an absolute continuous measure relative to the Borel - Lebesgue measure, significantly generalizing the earlier theorem by Lebesgue about the existence of an almost everywhere differentiable density.

Right after the appearance of Radon's paper, Maurice Fréchet noticed that Radon's result can be generalized for arbitrary measures, rather than Borel measures on $\mathcal{B}(\mathbb{R}^n)$. This led Nikodym to his 1930 generalization of Radon's theorem in the form very close to the present version. Consequently, a significant gap in integral theory existed between 1913 and 1930. Soon thereafter, in 1933, Nikodym's generalization led to the birth of measure-theoretic probability theory (in Andrey Kolmogorov's famous monograph, *Grundbegriffe der Wahrscheinlichkeitsrechnung*), the concept of conditional expectation, and an introduction to the theory of stochastic processes. Still, many consider Radon to be the father of the modern theory of integration.

Otto Nikodym, who is at the heart of one of the most significant results ever made in mathematics, was born on August 13, 1887, in eastern Poland, then belonging to the Russian empire. In 1919 he was among 16 mathematicians who founded the Polish Mathematical Society. Shortly after World War II, Nikodym's family moved to Belgium and then to France, where Nikodym was invited by the Institute of H. Poincaré to work on the mathematical foundations of quantum mechanics. (He published his results in numerous papers, and his monograph, *The Mathematical Apparatus for Quantum Theories*, was published by Springer-Verlag in 1966.) In 1948, he accepted a position in the United States at Kenyon College, Gambier, Ohio, where he stayed until his retirement. He died in 1974.

Johann Radon was born in Bohemia, Austria-Hungary, now Czech Republic, on December 16, 1887. He received his doctoral degree at the University of Vienna in 1910. Besides his best-known work on integration, Radon applied the calculus of variations to differential geometry that led to applications in number theory. He

discovered curves which are now called *Radon curves*. Next he worked on convex functions and sets and on the determination of functions from the values of their integrals on certain manifolds that contain the Radon transform and play an important role today, especially in medicine and geophysics. His wide interests also brought him to Riemannian geometry and geometrical problems that arose in the study of relativity.

Between 1919 and 1955 Radon had numerous appointments at Austrian and German universities. (In 1954 - 1955, he served as president of the University of Vienna.) In 1947, Radon became a full member of the Austrian Academy of Sciences. From 1948 to 1950, he was the president of the Austrian Mathematical Society. He died in Vienna in May 1956.

We introduce some preliminaries on the Radon - Nikodym Theorem (further embellished in Section 8).

5.6 Notation. Let $\mathfrak{M} = \mathfrak{M}(\Omega, \Sigma)$ be the set of all (positive) measures on (Ω, Σ). For a fixed measure $\mu \in \mathfrak{M}$, denote $\mathfrak{M}_\mu^{\ll} = \{\nu \in \mathfrak{M} : \nu \ll \mu\}$. (This set is not empty, because $\mu \in \mathfrak{M}_\mu^{\ll}$.) Define on $\mathcal{C}_+^{-1}(\Omega, \Sigma) = \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+))$ a mapping $\int_\mu$ such that for each $f \in \mathcal{C}_+^{-1}$,

$$\int_\mu f = \int_{(\cdot)} f d\mu = \nu(\,\cdot\,).$$

By Lemma 1.15 [*for* $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$, $\int f d\mu = 0$ *if and only if* $f \in [\mathbf{0}]_\mu^+$. *In particular, whenever* $\mu(A) = 0$, $\mathbf{1}_A f \in [\mathbf{0}]_\mu^+$ (*under the convention that* $0 \cdot \infty = 0$) $\Rightarrow \nu(A) = \int \mathbf{1}_A f d\mu = 0$], we have that for each measurable function $f \geq 0$ and for each measure μ on Σ, not only is $\int_{(\cdot)} f d\mu$ a measure, but it is also continuous with respect to μ. Thus,

$$\int_\mu : \mathcal{C}_+^{-1}(\Omega, \Sigma) \to \mathfrak{M}_\mu^{\ll}$$

is an into map. □

Now the Radon - Nikodym theorem asserts that if μ is σ-finite, for each $\nu \in \mathfrak{M}_\mu^{\ll}$, there exists a unique (up to the equivalence class

modulo μ) Radon - Nikodym density $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$ of ν relative to μ. This needs some clarification:

1) Given a function $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$, $\int_\mu f$ defines a measure that is absolutely continuous with respect to μ. As noticed above, this is done. Consequently, $[\mathcal{C}_+^{-1}(\Omega, \Sigma), \mathfrak{M}_\mu^{\ll}, \int_\mu]$ is an into map.

2) Recall [Definitions and Remarks 5.7(iii), Chapter 5] that the μ-almost everywhere property of equality of measurable functions generates an equivalence relation E_μ (the equivalence kernel of μ) on $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ and thus on $\mathcal{C}_+^{-1}(\Omega, \Sigma)$, as a subset of $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Consequently, $\mathcal{C}_+^{-1}(\Omega, \Sigma)|_\mu$ is a quotient subset of $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})|_\mu$ modulo E_μ, also referred to as *modulo* μ. On the other hand, by Corollary 1.20 [*If* μ *is* σ*-finite*, $f, g \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, *and* $\int_A f\, d\mu = \int_A g\, d\mu$, *for each* $A \in \Sigma$, *then* $f = g$ μ-a.e. *on* Ω], the mapping $\int_\mu$ generates E_μ as its equivalence kernel. Then, by Theorem 4.4, Chapter 1, there is a unique function

$$[\mathcal{C}_+^{-1}(\Omega, \Sigma)|_\mu = \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+)|_\mu, \mathfrak{M}_\mu^{\ll}, \boldsymbol{\int_\mu}],$$

such that

$$\int_\mu = \boldsymbol{\int_\mu} \circ \pi_{\mathrm{E}},$$

where π_{E} stands for the *projection of* $\mathcal{C}_+^{-1}(\Omega, \Sigma)$ *on its quotient* $\mathcal{C}_+^{-1}(\Omega, \Sigma)|_\mu$ *by* E_μ. (See Section 4, Chapter 1.)

Therefore, $\int_\mu$ literally turns to injective map $\boldsymbol{\int_\mu}$ (the reducer of $\int_\mu$) that now operates on the quotient set $\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+) \mid_\mu$.

3) The major claim (existence) of the Radon - Nikodym theorem is that the mapping $[\mathcal{C}_+^{-1}(\Omega, \Sigma)|_\mu, \mathfrak{M}_\mu^{\ll}, \boldsymbol{\int_\mu}]$ is also surjective. In other words, for each measure $\nu \in \mathfrak{M}_\mu^{\ll}$ (i.e., absolutely continuous with respect to μ), there is an equivalence class $[f]_\mu$ of Radon - Nikodym densities of ν relative to μ.

A compact version of the above arguments is as follows.

5.7 Theorem (Radon - Nikodym). *Let $\mu \in \mathfrak{M}(\Omega, \Sigma)$ be a σ-finite measure. Then $[\mathcal{C}_+^{-1}(\Omega, \Sigma)|_\mu\,, \mathfrak{M}_\mu^{\ll}, \int_{\boldsymbol{\mu}}]$ is a bijective map.*

As mentioned, the uniqueness of the Radon - Nikodym density class is due to Corollary 1.20. The rest of the proof of Theorem 5.7 (existence) is rendered in Section 8, for more general classes of signed measures. □

By Radon - Nikodym's Theorem, the map $\int_{\boldsymbol{\mu}}$ is therefore invertible and its inverse, denoted by symbol $\frac{d}{d\mu}$, is also a map valued in $\mathcal{C}_+^{-1}(\Omega, \Sigma)|_\mu$. Thus, for any $\nu \in \mathfrak{M}_\mu^{\ll}$, there is a nonempty equivalence class $[f]_\mu$ of Radon - Nikodym densities of ν relative to μ and, for a fixed $\nu \in \mathfrak{M}_\mu^{\ll}$, we write

$$\tfrac{d\nu}{d\mu} := \textstyle\int_{\boldsymbol{\mu}}^{-1}(\nu) = \int_\mu^*(\{\nu\})$$

and call it the Radon - Nikodym *derivative of measure ν relative to the measure μ.* It should be clear that, unlike a Radon - Nikodym density, the Radon - Nikodym derivative $\frac{d\nu}{d\mu}$ is a μ-equivalence class of all Radon - Nikodym densities with respect to measure ν.

5.8 Examples.

(i) Let Ω be an uncountable set; let $\Sigma = \{A \in \mathcal{P}(\Omega)$: either A or A^c is countable$\}$; and let $\nu(A) = 0$ if A is countable and $\nu(A) = \infty$ if A^c is countable. Let μ be a counting measure on Σ [i.e., let $\mu(A) = |A|$ if A is finite]; otherwise, $\mu(A) = \infty$. Because the only μ-null set available is $\emptyset$, it immediately follows that $\nu \ll \mu$. However, we show that ν cannot have a density relative to μ. Assume the opposite. Let $g \in \mathcal{C}_+^{-1}$ be such that $g \in \frac{d\nu}{d\mu}$. Then, for all $\omega \in \Omega$, we have that:

$$\begin{aligned} 0 = \nu(\{\omega\}) &= \int_{\{\omega\}} g(x)\,\mu(dx) \\ &= \int \mathbf{1}_{\{\omega\}}(x) g(x) \mu(dx) = \int \mathbf{1}_{\{\omega\}}(x) g(\omega) \mu(dx) \\ &= g(\omega) \int \mathbf{1}_{\{\omega\}}(x) \mu(dx) = g(\omega)\mu(\omega) = g(\omega). \end{aligned}$$

This implies that $g \equiv 0$, which, in turn, yields that $\nu \equiv 0$. This is a contradiction. (For a further discussion see Problem 5.3.)

(ii) Let ε_a be a point mass on $(\mathbb{R}^n, \mathcal{B})$. We are interested in whether $\varepsilon_a \ll \lambda$. Let $B = \{a\} \in \mathcal{B}$. Then, $\lambda(\{a\}) = 0$ whereas $\varepsilon_a(\{a\}) = 1$. Therefore, by the Radon - Nikodym theorem, ε_a does not have any density relative to λ.

(iii) Let λ be the Borel - Lebesgue measure on $(\mathbb{R}^n, \mathcal{B})$ and let ν be a probability measure on the same measurable space such that $\nu \ll \lambda$. Because λ is σ-finite, by the Radon - Nikodym theorem, there exists a density, $f \in \frac{d\nu}{d\lambda}$, called the *probability density of* ν, such that $\nu = \int f d\lambda$. Should ν be a probability distribution, say $\mathbb{P}X^*$, induced by a random variable $X \in \mathcal{C}^{-1}(\Omega, \Sigma)$ that is absolutely continuous with respect to the Borel - Lebesgue measure, then f is referred to as the *probability density function* (pdf) of X.

Let $X\colon (\Omega, \Sigma) \to (\mathbb{R}, \mathcal{B})$ be a random variable with $\mathbb{P}X^* = \int f(\alpha, \sigma^2)\, d\lambda$, where

$$f(\alpha, \sigma^2; x) \;=\; \tfrac{1}{\sqrt{2\pi\sigma^2}} \exp\Big\{ -\tfrac{(x-\alpha)^2}{2\sigma^2} \Big\},\; x \in \mathbb{R}.$$

Then, X is called a *normal random variable with parameters* (α, σ^2); the corresponding probability density function of X, $f(\alpha, \sigma^2)$ is called the *normal density.* □

We discuss some more results related to the Radon - Nikodym theorem.

5.9 Definition. Let μ_1 and μ_2 be two measures on a measurable space (Ω, Σ). μ_1 is said to be *singular relative to* μ_2 (or *orthogonal to* μ_2) if there is an $A \in \Sigma$ such that

$$\mu_1(A) = \mu_2(A^c) = 0.$$

In this case we write $\mu_1 \perp \mu_2$. □

It should be clear that the orthogonality relation is symmetric.

5.10 Examples.

(i) Let ε_a be a point mass on $(\mathbb{R}^n, \mathcal{B})$ with $a \in \mathbb{R}^n$, and let $A = \{a\}^c$. Then, $\lambda(\{a\}) = \varepsilon_a(A) = 0$. Therefore, $\varepsilon_a \perp \lambda$. In Example 5.10$(ii)$, it was shown that ε_a is not absolutely continuous relative to λ. Now we have established another relation between ε_a and λ.

(ii) Let $\mu = c\varepsilon_a + b\lambda$, where $c, b \in \mathbb{R}_+ \backslash \{0\}$. It is clear that $c\varepsilon_a \perp \lambda$ and $b\lambda \ll \lambda$, which implies that measure μ is a sum of continuous and singular components relative to λ. In the general case, if μ and ν are measures on (Ω, Σ) such that ν is σ-finite, there is a unique decomposition of $\nu = \nu_a + \nu_s$, where ν_a is an absolute continuous measure relative to μ and ν_s is a singular measure relative to μ. This fact is due to the well-known Lebesgue decomposition theorem discussed in [Dsh2]. □

PROBLEMS

5.1 Prove Proposition 5.1. *Let* (Ω, Σ, μ) *be a measure space and let* $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$. *Then, the set function* $\nu(A) = \int\limits_A f d\mu$ *is a measure on* Σ.

5.2 Let μ and ν be measures on (Ω, Σ) such that $\nu \leq \mu$ and suppose μ is σ-finite. Show that there exists a Radon-Nikodym density $f \in \frac{d\nu}{d\mu}$, such that $0 \leq f \leq 1$.

5.3 Let μ and ν be measures on (Ω, Σ) such that $\nu \ll \mu$ and let ν be finite and $g \in \frac{d\nu}{d\mu}$. Denote $A = \{\omega \in \Omega : g(\omega) \neq 0\}$. Show that the restriction of μ on $\Sigma \cap A$ is σ-finite. Give an example where $\mu(A)$ is not finite.

5.4 Let π_α be a Poisson measure on $(\mathbb{R}, \mathcal{B})$. Investigate whether π_α is absolutely continuous or singular relative to λ.

5.5 Let ρ_1, ρ_2, and μ be measures on (Ω, Σ) such that $\rho_1 \perp \mu$ and $\rho_2 \perp \mu$. Show that $\rho_1 + \rho_2 \perp \mu$.

5.6 Let $\rho_1 \ll \mu$ and $\rho_2 \perp \mu$. Show that $\rho_1 \perp \rho_2$.

5.7 Prove that $\rho \ll \mu$ and $\rho \perp \mu$ imply that $\rho \equiv 0$.

5.8 Let μ and ν be σ-finite measures on (Ω, Σ). Show that μ and ν possess densities f and g, respectively, relative to $\rho = \mu + \nu$.

5.9 Is orthogonality transitive?

5.10. Let $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+))$, μ be a positive finite measure, $\Sigma_0 \subseteq \Sigma$ be a sub-σ-algebra, and $\mu_0 = \mathrm{Res}_{\Sigma_0}\mu$. Suppose $\Sigma_0 = \{\Omega, \emptyset, A, B = A^c\}$ where A is a measurable set. Define

$$g = \frac{1}{\mu_0(A)} \int\limits_A f d\mu_0 \mathbf{1}_A + \frac{1}{\mu_0(B)} \int\limits_B f d\mu_0 \mathbf{1}_B.$$

Show that g is a Radon - Nikodym density of $\nu = \int f d\mu$ with respect to μ_0.

5.11 Let X be a nonnegative random variable on a probability space (Ω, Σ, P) such that $X = c$ P-a.s. (a constant). Is $PX^* \ll \lambda$? If it is, give a Radon - Nikodym density of PX^* w.r.t. λ.

5.12 Prove that $\mu \perp \rho$ if and only if $\mu \wedge \rho = \theta$.

NEW TERMS:

density, 422
measure generated by an integral, 422
indefinite integral, 422
Radon - Nikodym density, 422
absolutely continuous measure, 424
$\nu \ll \mu$, means ν is absolutely continuous with respect to μ, 424
continuous measure, 424
Nikodym, Otto, 425
Saks, Stanislav, 425
Stieltjes, Thomas, 425
Riemann - Stieltjes integral, 425
Riesz, Frigyes, 425
Radon, Johann, 425
Fréchet Maurice, 426
Kolmogorov, Andrey, 426
$\mathfrak{M}(\Omega, \Sigma)$, the set of all measures on (Ω, Σ), 427
$\mathfrak{M}_\mu^{\ll} = \{\nu \in \mathfrak{M} : \nu \ll \mu\}$, the set of all measures continuous relative to μ, 427
$\int_\mu$, mapping defined on $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+)$, 427
$[\mathcal{C}_+^{-1}(\Omega, \Sigma)|_\mu = \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+)|_\mu, \mathfrak{M}_\mu^{\ll}, \boldsymbol{\int_\mu}]$, bijective map between Radon - Nikodym derivatives and continuous measures, 428
Radon - Nikodym theorem, 429
$\frac{d}{d\mu} = \boldsymbol{\int_\mu}$, reducer of $\int_\mu$, 429
Radon - Nikodym density, 429
$\frac{d\nu}{d\mu} := \boldsymbol{\int_\mu}^{-1}(\nu) = \int_\mu^*(\{\nu\})$, Radon - Nikodym derivative, 429
probability density, 430
probability density function, 430
normal random variable, 430
$f(\alpha, \sigma^2; x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ - \frac{(x-\alpha)^2}{2\sigma^2} \right\}$, normal probability density function, 430
singularity of a measures, 430
orthogonal measures, 430

6. SIGNED MEASURES

The situation below is motivational for the study of a more general class of set functions than those we called "measures." Let (Ω, Σ, μ) be a measure space and let $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. Define the following set function on Σ.

$$\nu(A) = \int_A f d\mu = \mu_1(A) - \mu_2(A),$$

where

$$\mu_1(A) = \int_A f^+ d\mu \quad \text{and} \quad \mu_2(A) = \int_A f^- d\mu.$$

It is easy to show that the set function ν has all the properties of a measure except for being positive. [See Problem 6.1 for a more general case of $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$.] However, in the above decomposition $\nu = \mu_1 - \mu_2$, the set function ν is represented by the difference of two measures. We study this type of set function, which we call a *signed measure*. We give a formal definition below, without saying anything about a decomposition that follows.

6.1 Definition. Let (Ω, Σ) be a measurable space. A set function $\nu : \Sigma \to \overline{\mathbb{R}}$ is called a *signed measure* if:

a) $\nu(\emptyset) = 0$,

b) for each $A \in \Sigma$, the value of $\nu(A)$ is well defined, i.e., it is either finite or $+\infty$ or $-\infty$,

c) ν is σ-additive.

To tell signed measures from nonnegative measures, we refer to the latter as *positive measures*. $\mathfrak{S}(\Omega, \Sigma)$ denotes the set of all signed measures on the measurable space (Ω, Σ). □

6.2 Remark. If a sequence of pairwise disjoint measurable sets $\{A_n\}$ is such that its union is finite under ν [i.e., $|\nu(\Sigma_{n=1}^{\infty} A_n)| < \infty$], then σ-additivity of ν [i.e. $|\nu(\Sigma_{n=1}^{\infty} A_n)| = |\Sigma_{n=1}^{\infty} \nu(A_n)|$] implies that the series $\Sigma_{n=1}^{\infty} \nu(A_n)$ must be absolutely convergent.

This, as we know, does not need to hold for series in the general case. The reader is encouraged to explain it. (See Problem 6.3.) □

6.3 Definitions.

(i) The signed measure is called *finite* if its range is a subset of $\mathbb{R}$. Otherwise, is called *infinite*. The triple (Ω, Σ, ν) is called the *signed measure space*. According to the type of the signed measure, the signed measure space is referred to as *finite* or *infinite*. The signed measure ν is called *σ-finite* if Σ admits a countable measurable partition $\{\Omega_n\}$ of ν-finite sets.

(ii) Sometimes, we need a notion of a *finite set under ν* (or a *ν-finite set* or just a *finite set*). This refers to a measurable set A with $|\nu(A)| < \infty$.

(iii) A measurable set P is called *ν-positive* (or just *positive*) if $\nu(P \cap A) \in [0, \infty]$ for all $A \in \Sigma$. A measurable set N is called *ν-negative* (or just *negative*) if $\nu(N \cap A) \in [-\infty, 0]$ for all $A \in \Sigma$. Obviously, P (N) is positive (negative) if and only if for any measurable subset E of P (N), $\nu(E) \geq 0$ (≤ 0).

We denote by $\mathfrak{P} \subseteq \Sigma$ the family of all positive sets under ν and by $\mathfrak{N} \subseteq \Sigma$ the family of all negative sets under ν.

(iv) A set function $\nu\colon \Sigma \to \overline{\mathbb{R}}$ is called *continuous from below* if for every monotone nondecreasing sequence $\{A_n\}\uparrow \subseteq \Sigma$ it holds that

$$\lim_{n\to\infty} \nu(A_n) = \nu\Big(\bigcup_{n=1}^{\infty} A_n\Big).$$

If the above property holds for a particular monotone nondecreasing sequence $\{A_n\}$, then ν is called *continuous from below on* $\{A_n\}$.

(v) Let $\{A_n\}$ be a monotone nonincreasing sequence of sets from Σ at least one of which is ν-finite. A set function $\nu\colon \Sigma \to \overline{\mathbb{R}}$ is said to be *continuous from above on* $\{A_n\}$ if

$$\lim_{n\to\infty} \nu(A_n) = \nu\Big(\bigcap_{n=1}^{\infty} A_n\Big). \tag{6.3}$$

The set function ν is *continuous from above on* Σ, if (6.3) holds for every monotone nonincreasing sequence $\{A_n\}\downarrow \subseteq \Sigma$ that has at least one ν-finite set. In particular, if $\{A_n\}\downarrow \emptyset$, (6.3) reduces to

$$\lim_{n\to\infty} \nu(A_n) = 0$$

and this is referred to as *continuity from above at the empty set* or, shortly, $\emptyset$-*continuity* of ν.

(vi) Any signed measure on the Borel σ-algebra is called a *signed Borel measure*. In particular, a signed Borel measure on $(\mathbb{R}^n, \mathcal{B})$, finite on d_e-bounded Borel sets, is said to be a *signed Borel - Lebesgue - Stieltjes measure.* □

6.4 Example. Let $\Omega = [-1,1]$, $\Sigma = \mathcal{B}(\mathbb{R}) \cap [-1,1]$, and let us define ν on Σ as $\nu(A) = \lambda(A \cap [0,1]) - \lambda(A \cap [-1,0])$. Obviously, $\nu(\emptyset) = 0$ and

$$\nu\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \underbrace{\{\lambda(A_n \cap [0,1]) - \lambda(A_n \cap [-1,0])\}}_{\nu(A_n)}.$$

Thus, ν is σ-additive. Furthermore, ν is finite. □

6.5 Remarks.

(i) Notice that if A is a finite set under ν, then any finite or countable decomposition of A consists of finite subsets under ν, that is not as obvious as for positive measures. (See also Problem 6.6.)

(ii) Unlike any positive measure, a signed measure fails to be monotone. (See Problem 6.5.) Yet some weak form of monotonicity applies to signed measures. (See Problems 6.2 and 6.6.)

(iii) We observe two obvious and noteworthy properties of positive and negative sets. Let $P \in \mathfrak{P}$ and $N \in \mathfrak{N}$ (be positive and negative sets). Then,

(a) $\mathcal{P}(P) \cap \Sigma \subseteq \mathfrak{P}$ and $\mathcal{P}(N) \cap \Sigma \subseteq \mathfrak{N}$ (where $\mathcal{P}$ stands for the power set). Indeed, if $A \subseteq P$, then $\nu(A) \geq 0$. If $B \subseteq A$, then B is also a subset of P and thus is positive.

(b) If $P \in \mathfrak{P}$ and $N \in \mathfrak{N}$, then $P \cap N$ is a null set. Indeed, as a subset of both P and N, $P \cap N$ must be simultaneously positive and negative. □

We continue with a few more introductory properties of signed measures.

6.6 Proposition. *Let (Ω, Σ, ν) be a signed measure space. If A and B are measurable sets such that B is ν-finite and $A \subseteq B$, then A is also ν-finite and $\nu(B \backslash A) = \nu(B) - \nu(A)$.*

6.7 Proposition. *Any signed measure ν is continuous from below.*

6.8 Proposition. *Let (Ω, Σ, ν) be a signed measure space and let $\{A_n\}$ be a monotone nonincreasing sequence of measurable sets such that for some n, $\nu(A_n)$ is finite. Then signed measure ν is continuous from above on $\{A_n\}$.*

6.9 Proposition. *Let ν be a finitely additive set function on the measurable space (Ω, Σ) such that $\nu(\emptyset) = 0$. If ν is continuous from below, then it is σ-additive. If ν is finite, then the continuity from above implies that ν is σ-additive. If ν is finite and continuous at the empty set, then ν is σ-additive.*

(The proofs to Propositions 6.6 - 6.9 are left for exercises; see Problems 6.6 - 6.9.)

6.10 Lemma. *Let (Ω, Σ, ν) be a signed measure space and let $A \in \Sigma$ be such that $-\infty < \nu(A) < 0$. Then, there is a negative subset N of A such that $-\infty < \nu(N) \leq \nu(A)$.*

Proof. If A is negative, i.e. it does not contain at least one subset E with $\nu(E) > 0$, then the statement of the lemma is proved. Otherwise, assume that A has at least one such subset and let

$$s_1 = \sup\{\nu(C)\colon C \subseteq A\} \in (0, \infty].$$

Now, if $s_1 = \infty$, then there is a subset $A_1 \subseteq A$ such that $\nu(A) \geq 1$ because by Proposition 6.6, all subsets of A are of finite measure under ν. If, however, s_1 is finite, then there is a subset $A_1 \subseteq A$ such that $\nu(A_1) \geq \frac{1}{2} s_1$. Combining both cases, we conclude that there is

an $A_1 \subseteq A$ such that $\nu(A_1) \geq \min\{\frac{1}{2}s_1, 1\}$. Now, consider the set $A\backslash A_1$ which has the following property by Proposition 6.6.

$$\nu(A\backslash A_1) = \nu(A) - \nu(A_1) < \nu(A),$$

because $\nu(A) > 0$. If $A\backslash A_1$ is negative and thus it does not contain any subset E with $\nu(E) > 0$, then we are done with the proof, with $N = A\backslash A_1$. Otherwise, we apply to $A\backslash A_1$ the same procedure as to A arriving at a subset A_2 of $A\backslash A_1$ with $\nu(A_2) \geq \min\{\frac{1}{2}s_2, 1\}$ (where $s_2 = \sup\{\nu(C) : C \subseteq A\backslash A_1\}$) and seeing set $A\backslash(A_1 + A_2)$ with property

$$\nu(A\backslash(A_1 + A_2)) = \nu(A\backslash A_1) - \nu(A_2) < \nu(A\backslash A_1) < \nu(A).$$

If the process we are going through sees its nth step, we arrive at a subset $A_n \subseteq A\backslash\sum_{i=1}^{n-1} A_i$, with $\nu(A_n) \geq \min\{\frac{1}{2}s_n, 1\}$, where

$$s_n = \sup\{\nu(C) : C \subseteq A\backslash\textstyle\sum_{i=1}^{n-1} A_i\},$$

and if $A\backslash\sum_{i=1}^{n-1} A_i$ is a negative set, we stop the algorithm. However, if we never happen to stop the process, we finally have the set

$$N = A\backslash\textstyle\sum_{n=1}^{\infty} A_n$$

with the properties asserted in the lemma. Indeed:

1) $-\infty < \nu(N) = \nu(A) - \nu(\sum_{n=1}^{\infty} A_n) < \nu(A) < 0$

2) Because $\nu(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \nu(A_n) > 0$ and finite, $\nu(A_n) \to 0$, and $\nu(A_n) \geq \min\{s_n \frac{1}{2}, 1\}$ making $s_n \to 0$.

3) On the other hand, by the definition of s_n, for each subset $E_n \subseteq A\backslash\sum_{i=1}^{n-1} A_i$, $\nu(E_n) \leq s_n \to 0$. Hence, each subset $E \subseteq N$ must have $\nu(E) \leq \lim_{n\to\infty} s_n = 0$ making N negative. □

The following theorem states that there is an (essentially unique) decomposition of set Ω into a positive and a negative set relative to a

given signed measure ν. This decomposition, referred to as a *Hahn decomposition*, leads us to the forthcoming Jordan decomposition of ν into the difference of two positive measures mentioned in the beginning of this section. We start with the following.

6.11 Definition. Let (Ω, Σ) be a measure space. Given a signed measure ν on Σ, suppose Ω can be partitioned into two sets, Ω^+ and Ω^- such that $\Omega^+ = P$ is a ν-positive set and $\Omega^- = N$ is a ν-negative set and $\Omega = \Omega^+ + \Omega^-$. Any such decomposition of Ω is called a *Hahn decomposition of* Ω (*with respect to* ν or rather *induced by* ν), in notation (Ω^+, Ω^-) or most commonly (P, N). □

The existence of such a decomposition of Ω is established in the following theorem.

6.12 Theorem (The Hahn Decomposition Theorem). *Let* (Ω, Σ, ν) *be a signed measure space. Then there exists a Hahn decomposition* (P, N) *of* Ω *with respect to* ν.

Proof. We assume, without loss of generality, that ν does not take the value $-\infty$. Otherwise, we consider $-\nu$ instead. If Ø is the only negative set of ν, then for each $A \in \Sigma$, then $\nu(A) \geq 0$. Furthermore, if there were a set A such that $\nu(A) < 0$, then by Lemma 6.10, there is a nonempty negative subset of A. Therefore, $(\Omega, \emptyset)$ is the trivial Hahn decomposition and we are done with the proof.

Otherwise, suppose there is at least one negative set of ν and then let

$$I = \inf\{\nu(E) : E \in \Sigma \text{ and } E \text{ is } \nu\text{-negative}\}.$$

Clearly, $I \leq 0$. Then, there is a sequence $\{N_n\}$ of negative sets with $\lim_{n\to\infty} \nu(N_n) = I$. Because of Problem 6.10,

$$N := \bigcup_{n=1}^{\infty} N_n$$

is also a negative set. Regarding B_n as $\bigcup_{k=1}^{n} N_k$, we have $\{B_n\}$ as a monotone nondecreasing sequence of negative sets $\uparrow N$ and hence, by Proposition 6.7, $\lim_{n\to\infty}\nu(B_n) = \nu(N)$. Furthermore, because $B_n \backslash N_n \subseteq B_n$ and B_n is negative, $\nu(B_n\backslash N_n) \leq 0$ and $\nu(B_n\backslash N_n) = \nu(B_n) - \nu(N_n)$ thus implying that $\nu(B_n) \leq \nu(N_n)$. The latter yields that $\nu(N) \leq I$. On the other hand, as for a negative set, $\nu(N) \geq I$, and thus combining the two inequalities we have $\nu(N) = I$.

Now we show that $P = N^c$ is a ν-positive set. If this were not the case, then there is at least one measurable subset A of P with $-\infty < \nu(A) < 0$ and then, by Lemma 6.10, there is a measurable, negative subset B of A with $\nu(B) \leq \nu(A)$; in particular, $\nu(B) < 0$. Then, $B + N$ makes a negative set such that $\nu(B + N) = \nu(B) + \nu(N) < \nu(N) = I$, which contradicts the fact that I is the ν-limit inferior of all negative sets.

If $\nu \in [-\infty, \infty)$, as we mentioned in the beginning of the proof, we instead consider $-\nu \in (-\infty, \infty]$ with the corresponding Hahn decomposition (P, N). Then, the Hahn decomposition of Ω with respect to ν is obviously (N, P). □

6.13 Theorem (Uniqueness of the Hahn decomposition). *A Hahn decomposition is unique in the following sense. If there is another Hahn decomposition* (P', N') *then* $P\Delta P'$ *and* $N\Delta N'$ *are* ν*-null sets.* □

The uniqueness of the Hahn decomposition is left for an exercise (Problem 6.13).

6.14 Notation. Because by Theorem 6.13, Hahn decompositions form a single equivalence class modulo ν, we denote this class by symbol $[H^+, H^-]_\nu$ and call it the *Hahn class* (*modulo* ν). □

Although any Hahn decomposition is a decomposition of Ω (induced by a signed measure ν), the so-called "Jordan decomposition" in the forthcoming discussion is the one of the signed measure itself. It states that each signed measure can be uniquely represented by the difference of two positive measures, which is kind of a converse of Problem 6.4. We start with the following.

6.15 Definition. Let $(P, N) \in [H^+, H^-]_\nu$. Denote

$$\nu^+(\,\cdot\,) = \nu(P \cap \,\cdot\,)$$
$$\nu^-(\,\cdot\,) = -\nu(N \cap \,\cdot\,)$$
$$|\nu| = \nu^+ + \nu^-$$

and call all three of them the *Jordan variations*, of which ν^+ is the *positive*, ν^- is *negative*, and $|\nu|$ is *total the variation relative to* (P, N). □

6.16 Remarks.

(i) Obviously ν^+ and ν^- are positive measures and as the result, so is $|\nu|$.

(ii) $\nu = \nu^+ - \nu^-$, which can be readily shown.

(iii) As we show in Theorem 6.17, Jordan variations are invariant of any Hahn decomposition from $[H^+, H^-]_\nu$. □

6.17 Theorem. *Any two sets of Jordan variations relative to two Hahn decompositions from the same Hahn class* $[H^+, H^-]_\nu$ *are identical*.

Proof. Let $(P, N), (P', N') \in [H^+, H^-]_\nu$. To prove the theorem we need to show that for each $E \in \Sigma$,

$$\nu(P \cap E) = \nu(P' \cap E) \text{ and } \nu(N \cap E) = \nu(N' \cap E).$$

From $P = (P \cap P') + (P \backslash P')$ we have

$$P \cap E = (P \cap P' \cap E) + (P \backslash P') \cap E.$$

Because by the uniqueness of the Hahn decomposition $P \Delta P'$ is a null set, so is $P \backslash P'$ and so is $(P \backslash P') \cap E$. Thus,

$$\nu(P \cap P' \cap E) = \nu(P \cap E).$$

Interchanging the roles of P and P' we have

$$\nu(P \cap P' \cap E) = \nu(P' \cap E).$$

The same arguments yield that $\nu(N \cap E) = \nu(N' \cap E)$. □

6.18 Remarks and Examples.

(i) In light of Theorem 6.17 we need to modify Definition 6.16 about Jordan variations, in as much as they no longer depend upon any particular Hahn decomposition. In the sequel, we omit the phrase "relative to (P, N)" from Definition 6.16.

(ii) The pair (ν^+, ν^-) of positive and negative Jordan variations, as noticed restore $\nu = \nu^+ - \nu^-$. The latter representation is called the *Jordan decomposition* of a signed measure ν. According to Theorem 6.17, this decomposition is unique.

(iii) Again, by Theorem 6.12, there is a unique Hahn class $[H^+, H^-]_\nu$ induced by a signed measure ν. This class, in turn, generates a unique Jordan decomposition. We wonder if this process can be reversed in some sense. For instance, according to Problem 6.4, if μ_1 and μ_2 are any two positive measures on Σ, at least one of which is finite, then the difference $\nu = \mu_1 - \mu_2$ is a signed measure. Under what condition is $\mu_1 - \mu_2$ the Jordan decomposition of ν? (Here we are talking about a practical utility of existence Theorem 6.12.)

We point out one property of the Jordan decomposition $\nu = \nu^+ - \nu^-$, namely, that the positive and negative variations in the Jordan decomposition must be orthogonal in the same sense as that introduced in Section 5. So, we recall that two positive measures μ_1 and μ_2 are orthogonal if there is a measurable set A such that $\mu_1(A) = \mu_2(A^c) = 0$. In the event of the Jordan positive and negative variations any Hahn pair (P, N) can represent (A^c, A), namely, $\nu^+(N) = \nu^-(P = N^c) = 0$ and so they are "orthogonal with respect to (P, N)." Consequently, if we want to associate measures μ_1 and μ_2 above with a particular Hahn class $[H^+, H^-]_\nu$, we need to make sure that they are orthogonal ($\mu_1 \perp \mu_2$) with respect to any $(P, N) \in [H^+, H^-]_\nu$.

For example, let ε_a be a Dirac point mass and λ the Borel-Lebesgue measure on $\mathcal{B}(\mathbb{R})$. Then, as we know, $\varepsilon_a \perp \lambda$. For instance, $\lambda(\{a\}) = \varepsilon_a(\{a\}^c) = 0$. Then, if we define the signed measure ν as $\nu = \lambda - \varepsilon_a = \nu^+ - \nu^-$ (noticing that ε_a is finite), we

conclude that it is the Jordan decomposition. In this case, an associated Hahn pair is $(P = \{a\}^c, N = \{a\})$.

(iv) In another example, let π_α be a Poisson measure with parameter α defined as $\pi_\alpha = \sum_{n=0}^{\infty} e^{-\alpha} \frac{\alpha^n}{n!} \varepsilon_n$. The Poisson measure π_α is also orthogonal to λ. We can take $P = \mathbb{N}_0^c$ and see that $\lambda(P^c) = \pi_\alpha(P) = 0$. Thus, $\nu = \lambda - \pi_\alpha$ is the Jordan decomposition of ν relative to $(\mathbb{N}_0^c, \mathbb{N}_0)$. Now, there are other Hahn decompositions, besides (P, P^c) from the class $[H^+, H^-]_\nu$. For example, let positive and negative sets be $P' = \{-1, 0, 1, 2, \dots\}^c$ and $N' = \{-1, 0, 1, 2, \dots\}$, respectively. Or $\widetilde{P} = \mathbb{Z}^c$ and $\widetilde{N} = \mathbb{Z}$.

Yet another noteworthy Hahn pair is $\hat{N} = C \cup \mathbb{Z}$, where C is the Cantor set (defined in Example 3.11, Chapter 5), and $\hat{P} = \hat{N}^c$. Obviously, $\lambda\left(\hat{N}\right) = \pi_\alpha\left(\hat{P}\right) = 0$. Here $\hat{N}$, unlike other negative sets in the same context, is uncountable. □

In a nutshell, the above discussion can be formalized as follows.

6.19 Proposition. *Let μ_1 and μ_2 be two positive measures on a measurable space (Ω, Σ) such that $\mu_1 \perp \mu_2$ with respect to a decomposition $\Omega = \Omega^+ + \Omega^-$ [more specifically, $\mu_1(\Omega^-) = \mu_2(\Omega^+) = 0$]. Then, $\mu_1 - \mu_2 = \nu$ is a Jordan decomposition of ν relative to (Ω^+, Ω^-) which is an element of Hahn class $[H^+, H^-]_\nu$.*

Proof. Let $E^+ \subseteq \Omega^+ \cap \Sigma$. Because $\mu_2(\Omega^+) = 0$, by monotonicity of positive measures, $\mu_2(E^+) = 0$. Therefore, $\nu(E^+) = \mu_1(E^+) \geq 0$ and thus Ω^+ is a positive set relative to ν. Similarly, Ω^- is a negative set relative to ν. □

One of the very useful benefits of Proposition 6.19 arises in the integral representation of a signed measure that meets everyone's expectation that $\nu^+ = \int f^+ d\mu$, $\nu^- = \int f^- d\mu$ are positive and negative Jordan variations of the integral measure $\nu = \int f d\mu$. We give more details on this in the following.

6.20 Corollary. *Let (Ω, Σ, μ) be a positive measure space and let ν be the indefinite integral $\nu = \int f d\mu$ generated by $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ and μ. Then, ν is a signed measure and its Jordan variations are*:

$$\nu^+ = \int f^+ \, d\mu, \nu^- = \int f^- \, d\mu \quad \textit{and} \quad |\nu| = \int |f| d\mu,$$

where $P = \{f \geq 0\}$, $M = \{f \leq 0\}$, *and* $f^+ = f\mathbf{1}_P, f^- = f\mathbf{1}_M$. *Furthermore, a relevant Hahn pair* $(P, N) = (\Omega^+, \Omega^-) \in [H^+, H^-]_\nu$ *is* $P = \{f \geq 0\}$ *and* $N = \{f < 0\}$.

The proof of this proposition is quite simple and is left for the reader as Problem 6.18. □

Another significant property of the Hahn - Jordan decomposition is that it allows a more explicit representation of the Jordan variations playing a pivotal role in the forthcoming applications.

6.21 Proposition. *Let* (Ω, Σ, ν) *be a signed measure space. Then the positive, negative, and total variations of* ν *can be characterized as follows. Given any measurable set* $A \in \Sigma$,

(i) $\nu^+(A) = \sup\{\nu(E)\colon E \in \Sigma \cap A\}$.

(ii) $\nu^-(A) = \sup\{-\nu(E)\colon E \in \Sigma \cap A\} = -\inf\{\nu(E)\colon E \in \Sigma \cap A\}$.

(iii) $|\nu|(A) = \sup\{\sum_{k=1}^n |\nu(E_k)| : \{E_1, \ldots, E_n\} \subseteq \Sigma$ and $\sum_{k=1}^n E_k \subseteq A\}$.

Proof. Denote by (P, N) a Hahn decomposition of Ω with respect to ν and let

$$\nu_{\sup}(A) = \sup\{\nu(E)\colon E \in \Sigma \cap A\}$$

and

$$\nu_{\inf}(A) = \sup\{-\nu(E)\colon E \in \Sigma \cap A\}$$

$$= -\inf\{\nu(E)\colon E \in \Sigma \cap A\}.$$

(i) Clearly, $\nu^+(A) = \nu(A \cap P) \leq \nu_{\sup}(A)$. To prove the inverse inequality we notice that because $(P, \Sigma \cap P, \mathrm{Res}_{\Sigma \cap P}\nu)$ is a positive measure space, $\mathrm{Res}_{\Sigma \cap P}\nu$ is monotone and hence, for each $E \in \Sigma \cap A$,

$$\nu(E) = \nu(E \cap P) + \nu(E \cap N)$$

$$\leq \nu(E \cap P) \leq \nu(A \cap P) = \nu^+(A).$$

This yields the required inverse inequality

$$\nu_{\text{sup}}(A) \leq \nu^+(A)$$

and thereby proves part (i) of the proposition.

(ii) By interchanging the roles of P and N for $-\nu$, we have

$$(-\nu)^+(A) = -\nu(A \cap N) = \nu^-(A)$$

and therefore $\nu^- = -\nu_{\text{inf}}$.

(iii) Let $(P, N) \in [H^+, H^-]_\nu$ and let $A \in \Sigma$. Then,

$$|\nu|(A) = \nu(A \cap P) - \nu(N \cap A) = |\nu(E_1)| + |\nu(E_2)|,$$

where E_1, E_2 is a partition of A. Obviously,

$$|\nu|(A) \leq \sup\{\textstyle\sum_{k=1}^n |\nu(E_k)| : \{E_1, \ldots, E_n\} \subseteq \Sigma \text{ and}$$

$$\textstyle\sum_{k=1}^n E_k \subseteq A\} = : \alpha_A. \tag{6.21}$$

Let $E_1, \ldots, E_n$ be a disjoint measurable "subpartition" of A (i.e., a partition of a measurable subset of A). Then,

$$\textstyle\sum_{i=1}^n |\nu(E_i)| = \sum_{i=1}^n |\nu^+(E_i) - \nu^-(E_i)|$$

$$\leq \textstyle\sum_{i=1}^n |\nu^+(E_i) + \nu^-(E_i)| = \sum_{i=1}^n |\nu|(E_i)$$

$$= |\nu|(\textstyle\sum_{i=1}^n E_i) \leq |\nu|(A).$$

The latter inequality also holds for $\sup\{\sum_{i=1}^n |\nu(E_i)|\}$ and thus,

$$\alpha_A \leq |\nu|(A). \tag{6.21a}$$

Inequalities (6.21) and (6.21a) imply that $\alpha_A = |\nu|(A)$. □

6.22 Remarks.

(i) To summarize the Hahn - Jordan decomposition, by Proposition 6.21, we have that

$$\nu^+(A) = \sup\{\nu(E)\colon E \in \Sigma \cap A\} = \nu(A \cap P),$$

$$-\nu^-(A) = \inf\{\nu(E)\colon E \in \Sigma \cap A\} = \nu(A \cap N)$$

and

$$\nu(A) = \nu^+(A) - \nu^-(A).$$

These equations offer an obvious interpretation. Given a measurable set A, the signed measure ν attains its maximum and minimum values on two measurable disjoint subsets of A: $A \cap P$ and $A \cap N$, respectively; and the entire measure of A is the sum of these two values. In particular, it follows that P and N are the *ν-maximal* and *ν-minimal* subsets of Ω (in notation, $P = S$ and $N = I$) on which ν attains maximum and minimum values, respectively. This is due to the fact that $(P, \Sigma \cap P,\ \mathrm{Res}_{\Sigma\cap P}\nu)$ is a positive measure space and hence $\mathrm{Res}_{\Sigma\cap P}\nu$ is monotone. A similar argument explains why ν attains a minimum value on N.

(ii) Let (Ω, Σ, ν) be a signed measure space. If ν is a positive measure, then obviously, $S = \Omega$ and $I = \emptyset$, where S and I are introduced in (i). Consider the case with $\nu = \int f d\rho$, where ρ is a positive measure on (Ω, Σ) and $f \in L^1(\Omega, \Sigma, \rho)$. Then,

$$\nu(A) = \int_A f d\rho = \int_{A\cap\{f\geq 0\}} f d\rho + \int_{A\cap\{f<0\}} f d\rho.$$

Therefore,

$$\int_{\{f<0\}} f d\rho \leq \nu(A) \leq \int_{\{f\geq 0\}} f d\rho, \forall\ A \in \Sigma, \tag{6.22}$$

and consequently, $S = \{f \geq 0\}$ and $I = \{f < 0\}$.

Inequality (6.22) is trivial in the sense that $\nu(A) = \nu^+(A) - \nu^-(A) \leq \nu^+(A)$ and $\nu(A) \geq -\nu^-(A)$. All we need is to recognize the left- and the right-hand sides of (6.22) as the negative and positive Jordan variations of ν. Corollary 6.20 helps interpret this context of (6.22). Indeed, we have that

$$\nu(A) \leq \nu^+(A) = \nu(S \cap A) = \int f^+ \mathbf{1}_A d\rho$$

and that $\nu^-(A) = -\nu\{I \cap A\} = -\int f^- \mathbf{1}_A d\rho \geq -\nu(A)$.

(iii) If μ and ρ are two positive measures at least one of which is finite, the difference $\nu = \mu - \rho$ is a signed measure. However, it is not the Jordan decomposition of ν unless μ and ρ are orthogonal (as we learned from Proposition 6.19). Let (Ω, Σ, ν) be a signed measure space. Denote by $\nu_E = \text{Res}_{\Sigma \cap E}\nu$, where E is a measurable set. To obtain the Jordan decomposition of $\nu = \mu - \rho$, we need any Hahn decomposition of Ω with respect to ν. Say, (P, N) is one. Then, Jordan variations are

$$\nu^+ = \nu_P = \mu_P - \rho_P \tag{6.22a}$$

and

$$\nu^- = -\nu_N = \rho_N - \mu_N. \tag{6.22b}$$

We can also make use of formulas of (i) and (ii) of Proposition 6.21 to determine the positive and negative variations. □

6.23 Examples.

(i) Let ε_0 be the point mass and $\mathbb{P}$ a probability measure on $(\mathbb{R}, \mathcal{B})$. We find a Hahn decomposition of the signed measure $\nu = \mathbb{P} - \varepsilon_0$. We show that $I = \{0\}$ is a ν-minimal (and negative) set in light of Remark 6.22(i). For an $A \in \mathcal{B}$,

$$\nu(A \cap I^c) = \mathbb{P}(A \cap I^c) \geq 0,$$

and either

$$\nu(A \cap I) = \mathbb{P}(\{0\}) - \varepsilon_0(\{0\}) = \mathbb{P}(\{0\}) - 1, \text{ with } 0 \in A$$

or

$$\nu(A \cap I) = 0, \text{ with } 0 \notin A,$$

which implies that $\nu(A \cap I) \leq 0$. Using relations (6.22a) and (6.22b), we have the Jordan decomposition of ν:

$$\nu^+(A) = \nu(A \cap I^c) = \mathbb{P}(A \cap I^c),$$

and

$$\nu^-(A) = \mathbf{1}_A(0)(1 - \mathbb{P}(\{0\}).$$

Note that $\{0\} = I$ is the set where ν attains its minimum.

(ii) Let $\nu = \lambda - \mu$, where λ is the Lebesgue measure on $(\mathbb{R}, \mathcal{B})$ and μ is the *geometric* measure defined as

$$\mu = \textstyle\sum_{n=1}^{\infty} pq^{n-1}\varepsilon_n, \, 0 < p < 1, \, q = 1 - p.$$

Clearly, $N = \{1, 2, \dots\}$ is a negative set relative to ν, whereas $P = N^c$ is a positive set. Thus, (P, N) is a Hahn decomposition of $\mathbb{R}$ relative to ν and, consequently, for every Borel set A,

$$\nu^+(A) = (\lambda - \mu)(A \cap \{1, 2, \dots\}^c)$$

and

$$\nu^-(A) = (\mu - \lambda)(A \cap \{1, 2, \dots\})$$

represent the Jordan decomposition of ν. Because N is a λ-null set, the latter reduces to

$$\nu^-(A) = \mu(A \cap \{1, 2, \dots\}).$$

Therefore, ν attains its minimum at N and its value is -1, whereas the maximum value of ν is ∞ and it is attained at N^c. □

6.24 Remark (Signed Measures as a Lattice).

As we worked on positive measures in Chapter 5, Section 1, imposing in $\mathfrak{M}(\Omega, \Sigma)$ some space structure, we similarly approach $\mathfrak{S}(\Omega, \Sigma)$ as the set of all signed measures on a measurable space (Ω, Σ). It cannot be a linear space to begin with, in as much as the

sum of two signed measures need not be a signed measure, for obvious reasons that $\infty - \infty$ (we can encounter) is not defined.

However, we can structure a smaller set $\mathfrak{S}_*(\Omega, \Sigma)$ of finite signed measures by making it a vector space over field $\mathbb{R}$ and including θ as the zero measure. It can be easily shown (Problem 6.23) that, unlike its predecessor $\mathfrak{M}(\Omega, \Sigma)$, $(\mathfrak{S}_*(\Omega, \Sigma), \mathbb{R}, +, \cdot)$ is indeed a vector space. Furthermore, in Chapter 1, Section 1, we showed that $(\mathfrak{M}(\Omega, \Sigma), \preceq)$ was a lattice. We prove that similar lattice operations establish $(\mathfrak{S}_*(\Omega, \Sigma), \preceq)$ also as a lattice and not only that. □

In the next two lemmas whose proofs we leave to the reader, we demonstrate that $(\mathfrak{S}_*(\Omega, \Sigma), \preceq)$ is indeed a lattice and further enhance it to a vector lattice associating the Jordan variations with the familiar lattice components including $\nu^+, \nu^-,$ and modulus $|\nu|$.

6.25 Lemma. *Let $\nu_1, \nu_2 \in \mathfrak{S}_*(\Omega, \Sigma)$. Then the set function*

$$\rho(A) = \sup\{\nu_1(E) + \nu_2(A \backslash E) : E \in \Sigma \cap A\}$$

is a finite signed measure on Σ.

(See Problem 6.19.) □

6.26 Lemma. *Let $\nu_1, \nu_2 \in \mathfrak{S}_*(\Omega, \Sigma)$. Then the set function*

$$\sigma(A) = \inf\{\nu_1(E) + \nu_2(A \backslash E) : E \in \Sigma \cap A\}$$

is a finite signed measure on Σ.

(See Problem 6.20.) □

6.27 Remark. We introduce on $\mathfrak{S}_*(\Omega, \Sigma)$ the partial order $\preceq$ such that $\nu \preceq \nu$ if and only if $\nu(A) \leq \nu(A)$ for all $A \in \Sigma$. □

The theorem below states that $(\mathfrak{S}_*(\Omega, \Sigma), \preceq)$ is a lattice.

6.28 Theorem. *Let ρ and σ be defined as in Lemmas* 6.25 - 6.26. *Then, $\rho = \nu_1 \vee \nu_2$ and $\sigma = \nu_1 \wedge \nu_2$, and consequently, the partially ordered set $(\mathfrak{S}_*(\Omega, \Sigma), \preceq)$ is a lattice.*

Proof. (i) Let $\mathcal{U}(\{\nu_1, \nu_2\}) \subseteq \mathfrak{S}_*(\Omega, \Sigma)$ denote the set of all upper bounds of the two-element set $\{\nu_1, \nu_2\}$. We show that $\rho \in \mathcal{U}(\{\nu_1, \nu_2\})$. In other words, we show that for each $A \in \Sigma$,

$\rho(A) \geq \nu_1(A)$ and $\rho(A) \geq \nu_2(A)$, in notation, $\rho \succeq \nu_1$ and $\rho \succeq \nu_2$. Indeed, for any measurable subset $E \subseteq A$,

$$\rho(A) \geq \nu_1(E) + \nu_2(A\backslash E).$$

Choosing $E = A$ or $E = \emptyset$, we arrive at $\rho \succeq \{\nu_1, \nu_2\}$ and thus $\rho \in \mathcal{U}(\{\nu_1, \nu_2\})$. Similarly, it can be shown that the measure σ from Lemma 6.24 belongs to the set $\mathcal{L}(\{\nu_1, \nu_2\})$ of all lower bounds of $\{\nu_1, \nu_2\}$, in notation, $\sigma \preceq \{\nu_1, \nu_2\}$.

(ii) Let $u \in \mathcal{U}(\{\nu_1, \nu_2\})$ be another upper bound of the set $\{\nu_1, \nu_2\}$. We show that $\rho \preceq u$, which validates that ρ is the least upper bound of $\{\nu_1, \nu_2\}$. Let $E \subseteq A$ be two measurable sets. Then,

$$\nu_1(E) + \nu_2(A\backslash E) \leq u(E) + u(A\backslash E) = u(A)$$

implying that $\rho \preceq u$ and that $\rho = \nu_1 \vee \nu_2$.

Now, let $l \in \mathcal{L}(\{\nu_1, \nu_2\})$. Similarly,

$$l(A) = l(E) + l(A\backslash E) \leq \nu_1(E) + \nu_2(A\backslash E) \tag{6.28}$$

holding for all $E \subseteq A$. Therefore, the inequality sustains for the infimum on the right of (6.27). Consequently, $l \preceq \sigma$ and thus σ is the greatest lower bound of $\{\nu_1, \nu_2\}$. □

6.29 Remark and Notation (Signed Measures as a Riesz Space). As $\mathfrak{S}_*(\Omega, \Sigma)$ is authenticated as a vector lattice (or synonymously, *Riesz space*), we use the notation $(\mathfrak{S}_*(\Omega, \Sigma), \mathbb{R}, +, \cdot, \preceq)$ as for the vector lattice carried by the set $\mathfrak{S}_*(\Omega, \Sigma)$ of all finite signed measures.

Furthermore, from Lemmas 6.25 and 6.26 and Theorem 6.28 we concluded that ρ and σ defined as

$$\rho(A) = \sup\{\nu_1(E) + \nu_2(A\backslash E) : E \in \Sigma \cap A\}$$
$$\sigma(A) = \inf\{\nu_1(E) + \nu_2(A\backslash E) : E \in \Sigma \cap A\}$$

were $\rho = \nu_1 \vee \nu_2$ and $\sigma = \nu_1 \wedge \nu_2$. Compared to Proposition 6.21 $(i\text{-}ii)$, we see that $\nu \vee \theta = \nu^+$ and $\nu \wedge \theta = \nu^-$, whereas $|\nu| = \nu^+$

$+ \nu^-$, as per Definition 6.15 (called there the *Jordan variations*), are consistent with the associated lattice operations.

Along this side, we can get an access to the whole reserve of properties made available for the general lattice operations from Chapter 1, Section 7, which no doubt is a useful asset in studying space $(\mathfrak{S}_*(\Omega, \Sigma), \mathbb{R}, +, \cdot, \preceq)$. □

6.30 Definition. We further embellish the space structure of $(\mathfrak{S}_*(\Omega, \Sigma), \mathbb{R}, +, \cdot, \preceq)$ by introducing a real nonnegative function on $\mathfrak{S}_*(\Omega, \Sigma)$ by

$$\|\nu\| = |\nu|(\Omega),$$

which we claim to be a norm. □

6.31 Proposition (The Norm Properties).

(i) $\| \cdot \| \geq 0$.

(ii) $\|\nu\| = 0$ if and only if $\nu = \theta$.

(iii) $\|\alpha\nu\| = |\alpha|\|\nu\|$.

(iv) $\|\nu_1 + \nu_2\| \leq \|\nu_1\| + \|\nu_2\|$.

Proof. In all instances below we make use of properties of lattice operations from Chapter 1, Section 7.

(i) Is obvious. Furthermore, by Problem 6.17, $\| \cdot \|$ is finite.

(ii) $0 = |\nu|(\Omega) = \nu^+(\Omega) + \nu^-(\Omega) = 0$ implies that $\nu = \theta$.

(iii) $\|\alpha\nu\| = |\alpha\nu|(\Omega) = |\alpha||\nu|(\Omega)$, due to Problem 7.17$(vi)$.

(iv) Due to the triangle inequality for modulus $|\nu_1 + \nu_2| \preceq |\nu_1| + |\nu_2|$ [Problem 7.18(v)], we have $\|\nu_1 + \nu_2\| = |\nu_1 + \nu_2|(\Omega) \leq |\nu_1|(\Omega) + \nu_2(\Omega)$. □

The introduction of a norm on $\mathfrak{S}_*(\Omega, \Sigma)$ is fully rewarded in the following theorem.

6.32 Theorem. *The normed vector space* $(\mathfrak{S}_*(\Omega, \Sigma), \| \cdot \|)$ *is Banach*.

Proof. Let $\{\nu_n\} \subseteq \mathfrak{S}_*(\Omega, \Sigma)$ be a Cauchy sequence. Then, for each $\varepsilon > 0$, there is an $N(\varepsilon)$-tail of $\{\nu_n\}$. Thus we have

$$\begin{aligned} \varepsilon > \|\nu_m - \nu_n\| &= |\nu_m - \nu_n|(\Omega) \\ &\geq |\nu_m - \nu_n|(A) \geq |\nu_m(A) - \nu_n(A)|, \end{aligned} \tag{6.32}$$

for each $A \in \Sigma$, implying that $\{\nu_n\}$ is a Cauchy sequence uniformly on Σ. Therefore, for each $A \in \Sigma$, $\{\nu_n(A)\}$ is a Cauchy sequence of real numbers and as such, it is convergent to a finite real number, say $\nu(A)$. Therefore, ν is a set function on Σ, to which the sequence $\{\nu_n\}$ converges setwise on Σ.

Furthermore, by fixing n in (6.32) and passing to the limit in m we get

$$|\nu(A) - \nu_n(A)| \leq \varepsilon, n \geq N \text{ and for all } A \in \Sigma. \tag{6.32a}$$

We show that (i) ν is a measure and (ii) $\|\nu - \nu_n\| \to 0$.

(i) Obviously, $\nu(\emptyset) = 0$. Now, let $A_1, \ldots, A_k$ be disjoint measurable sets. Then,

$$\begin{aligned} \nu\Big(\sum_{i=1}^k A_i\Big) &= \lim_{n\to\infty} \nu_n\Big(\sum_{i=1}^k A_i\Big) = \sum_{i=1}^k \lim_{n\to\infty} \nu_n(A_k) \\ &= \sum_{i=1}^k \nu(A_k). \end{aligned}$$

Therefore, ν is finitely additive.

Next, we show that ν is continuous at the empty set. Let $\{A_k\} \downarrow \subseteq \Sigma$ such that $\bigcap_{k=1}^{\infty} A_k = \emptyset$. Due to the uniform convergence of $\{\nu_n\}$ on Σ to ν, as per (6.32a), for each $\varepsilon > 0$, there is an $N(\varepsilon)$-tail of $\{\nu_n\}$ such that $|\nu_n(A) - \nu(A)| \leq \varepsilon, n \geq N$. Now, by continuity at $\emptyset$, for each ν_n, we have $\lim_{k\to\infty} \nu_n(A_k) = 0$. In other words, for each $\varepsilon > 0$, there is an M such that for all $k \geq M$, $|\nu_n(A_k)| < \frac{\varepsilon}{2}$. then, for all $k \geq M$ and $n \geq N$,

$$|\nu(A_k)| \leq |\nu(A_k) - \nu_n(A_k)| + |\nu_n(A_k)| < \varepsilon.$$

This implies that ν is Ø-continuous. By Proposition 6.9, then ν is σ-additive, as it meets all assumptions of Proposition 6.9, namely, ν is finite, finitely additive, and continuous at the empty set.

(ii) We use Proposition 6.21(i, ii) according to which

$$(\nu_n - \nu)^+(\Omega) = \sup\{(\nu_n - \nu)(E) : E \in \Sigma\}$$

and

$$(\nu_n - \nu)^-(\Omega) = \sup\{-(\nu_n - \nu)(E) : E \in \Sigma\}$$

considering that $\nu_n - \nu \in \mathfrak{S}_*(\Omega, \Sigma)$. Furthermore, using (6.32a),

$$(\nu_n - \nu)(A) \leq |(\nu_n - \nu)(A)| \leq \varepsilon$$

as well as

$$-(\nu_n - \nu)(A) \leq |(\nu_n - \nu)(A)| \leq \varepsilon$$

both holding for $n \geq N$ and uniformly for all $A \in \Sigma$. They in turn imply that

$$(\nu_n - \nu)^+(\Omega) + (\nu_n - \nu)^-(\Omega) = \|\nu_n - \nu\| \leq 2\varepsilon,$$

good for all $n \geq N$ and thereby the claimed convergence $\nu_n \to \nu$ in the norm.

In conclusion, $(\mathfrak{S}_*(\Omega, \Sigma), \|\cdot\|)$ is a complete NVS. □

6.33 Remark. The space $(\mathfrak{S}_*(\Omega, \Sigma), \mathbb{R}, +, \cdot, \preceq, \|\cdot\|)$ turns out to be richly endowed with a variety of structures and operations making it an awesome paradigm of an NVS, which is a lattice and Banach. We call it for short a *Banach lattice*. □

PROBLEMS

6.1 Show that the integral set function ν introduced in the beginning of this section for the case of an $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ is a signed measure (i.e., $\nu(\emptyset) = 0$) for each $A \in \Sigma, \nu(A)$ is either finite or $+\infty$ or $-\infty$, and ν is σ-additive on Σ.

6.2 Let ν be a signed measure on a measurable space (Ω, Σ) and let A be a finite measurable set. Show that if $A = A_1 + A_2$ is any decomposition of A, then either A_1 and A_2 are finite. Furthermore, show that if $\nu(A) = \infty$, then at least one of the subsets A_1 or A_2 must be positively infinite. If $A = \Sigma_{n=1}^{\infty} A_n$ is a countable measurable decomposition of a finite set A, then all sets A_n's are finite. In particular, it follows that any subset of a finite set must be finite.

6.3 Why must the series $\Sigma_{n=1}^{\infty} \nu(A_n)$ in Remark 6.2 be absolutely convergent?

6.4 Let μ_1 and μ_2 be two positive measures on a measurable space (Ω, Σ) at least one of which is finite. Show that $\nu = \mu_1 - \mu_2$ is a signed measure.

6.5 Show by a counterexample that a signed measure is not monotone.

6.6 Prove Proposition 6.6. *If A and B are measurable sets such that B is ν-finite and $A \subseteq B$, then A is also ν-finite and $\nu(B \backslash A) = \nu(B) - \nu(A)$.*

6.7 Prove Proposition 6.7. *Any signed measure ν is continuous from below.*

6.8 Prove Proposition 6.8. *Let (Ω, Σ, ν) be a signed measure space and $\{A_n\}$ be a monotone nonincreasing sequence of measurable sets such that for some $n \in \mathbb{N}$, $\nu(A_n)$ is finite. Then the signed measure ν is continuous from above on $\{A_n\}$.*

6.9 Prove Proposition 6.9. *Let ν be a finitely additive set function on the measurable space (Ω, Σ) such that $\nu(\emptyset) = 0$. If ν is continuous from below, then it is σ-additive. If ν is finite, then the continuity from above implies that ν is σ-additive. If ν is finite and continuous at the empty set, then ν is σ-additive.*

6.10 Let $\mathfrak{P}$ and $\mathfrak{N}$ be the families of positive and negative sets, respectively. Show that $\mathfrak{P}$ and $\mathfrak{N}$ contain the empty set and they are closed with respect to differences and at most countable unions and intersections.

6.11 Let $\mathfrak{P}$ and $\mathfrak{N}$ be the families of positive and negative sets, respectively. Show that ν is monotone on either family.

6.12 Let (Ω, Σ, ν) be a signed measure space and let $A \in \Sigma$ be such that $0 < \nu(A) < \infty$. Show that there is a positive subset P of A, with $\nu(A) \leq \nu(P) < \infty$. (*Hint*: Use a similar approach as in the proof of Lemma 6.10.)

6.13 Show that all Hahn decompositions form a unique equivalence class.

6.14 Let ν be a signed measure on (Ω, Σ) represented by a difference of two positive measures $\nu = \mu_1 - \mu_2$. Show that $\mu_1 \geq \nu^+$ and $\mu_2 \geq \nu^-$.

6.15 Show that any signed Borel - Lebesgue - Stieltjes measure on $(\mathbb{R}^n, \mathcal{B})$ is σ-finite.

6.16 Let (Ω, Σ, ν) be a signed measure space. Show that $|\nu|(A) = 0$, for an $A \in \Sigma$, if and only if $\nu(S) = 0$ for each $S \in \Sigma \cap A$. Show by example that $\nu(A) = 0$ is not sufficient for $|\nu|(A) = 0$.

6.17 Let $\nu \in \mathfrak{S}_*(\Omega, \Sigma)$; that is, ν is a finite signed measure. Is it true that ν is bounded?

6.18 Prove Corollary 6.20. *Let (Ω, Σ, μ) be a positive measure space and let ν be the indefinite integral $\nu = \int f d\mu$ generated by $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ and μ. Then, ν is a signed measure and its Jordan variations are*:

$$\nu^+ = \int f^+ \, d\mu, \nu^- = \int f^- \, d\mu \text{ and } |\nu| = \int |f| d\mu,$$

where $P = \{f \geq 0\}$, $M = \{f \leq 0\}$, *and* $f^+ = f\mathbf{1}_P, f^- = f\mathbf{1}_M$. *Furthermore, a relevant Hahn pair* $(P, N) = (\Omega^+, \Omega^-) \in [H^+, H^-]_\nu$ *is* $P = \{f \geq 0\}$ *and* $N = \{f < 0\}$.

6.19 Prove Lemma 6.25. *Let* $\nu_1, \nu_2 \in \mathfrak{S}_*(\Omega, \Sigma)$. *Then the set function*

$$\rho(A) = \sup\{\nu_1(E) + \nu_2(A\backslash E) : E \in \Sigma \cap A\}$$

is a finite signed measure on Σ.

6.20 Prove Lemma 6.26. *Let* $\nu_1, \nu_2 \in \mathfrak{S}_*(\Omega, \Sigma)$. *Then the set function*

$$\sigma(A) = \inf\{\nu_1(E) + \nu_2(A\backslash E) : E \in \Sigma \cap A\}$$

is a finite signed measure on Σ.

6.21 Let $\nu, \rho \in \mathfrak{S}_*(\Omega, \Sigma)$. Show that $\nu + \rho \in \mathfrak{S}_*(\Omega, \Sigma)$. Is it true that

$$\begin{aligned} \nu + \rho &= (\nu + \rho)^+ - (\nu + \rho)^- \\ &= \nu^+ + \rho^+ - \nu^- - \rho^-? \end{aligned}$$

6.22 Let $\nu \in \mathfrak{S}(\Omega, \Sigma)$ and let $a \in \mathbb{R}$. Show that $a\nu \in \mathfrak{S}(\Omega, \Sigma)$. Is it true that $a\nu = (a\nu)^+ - (a\nu)^- = a\nu^+ - a\nu^-$? Will there be any changes in the Hahn class?

6.23 Show that $\mathfrak{S}_*(\Omega, \Sigma)$ is a vector space over the field $\mathbb{R}$.

NEW TERMS:

signed measure, 434
positive measure, 434
$\mathfrak{S}(\Omega, \Sigma)$, set of all signed measures on (Ω, Σ), 434
finite and infinite signed measures, 435
σ-finite signed measure, 435
finite (ν-finite) set, 435
positive (ν-positive) set, 435
negative (ν-negative) set, 435
$\mathfrak{P}$, $\mathfrak{N}$, families of positive and negative sets, 435
continuous from below set functions, 435
continuous from above set functions, 435
continuity from above at the empty set, 435
$\emptyset$-continuity of a signed measure, 436
signed Borel measure, 436
signed Borel - Lebesgue - Stieltjes measure, 436
Hahn decomposition, 439
Hahn decomposition theorem, 439
$[H^+, H^-]_\nu$, Hahn class modulo ν, 440
$\nu^+(\,\cdot\,)$, $\nu^-(\,\cdot\,)$, $|\nu| = \nu^+ + \nu^-$, Jordan variations, 441
positive and negative variations of a signed measure, 441
total variation of a signed measure, 441
Jordan decomposition of a signed measure, 442
signed measure generated by the integral, 443
characterization of Jordan variations, 444
ν-maximal and ν-minimal sets, 446
geometric measure, 448
signed measures as a lattice, 448
$\mathfrak{S}_*(\Omega, \Sigma)$, set of all finite signed measures, 449
$(\mathfrak{S}_*(\Omega, \Sigma), \preceq)$, vector lattice of finite signed measures, 449
signed measures as a Riesz space, 450
$\|\nu\| = |\nu|(\Omega)$, norm of signed measure ν, 451
$(\mathfrak{S}_*(\Omega, \Sigma), \|\cdot\|)$, normed vector space of finite signed measures, 451
$((\mathfrak{S}_*(\Omega, \Sigma), \mathbb{R}, +, \cdot, \preceq, \|\cdot\|)$, Banach lattice of finite signed measures, 453

7. COMPLEX MEASURES AND INTEGRATION

Our next embellishment of the space $\mathfrak{S}_*(\Omega, \Sigma)$ of finite signed measures is the space of all complex measure which plays a significant role in integration.

7.1 Definitions.

(i) Let (Ω, Σ) be a measurable space. A set function ξ on Σ is said to be a *complex measure* if:

(a) ξ is valued in $\mathbb{C}$. (Notice that being valued in $\mathbb{C}$, ξ must not have infinite values and, therefore, of those signed or positive measures only finite ones can be qualified as complex measures.)

(b) $\xi(\emptyset) = 0$.

(c) ξ is σ-additive. Analogous to the signed measures [see Remark 6.2(ii)], equation

$$| \xi(\Sigma_{n=1}^{\infty} A_n) | \ = \ | \Sigma_{n=1}^{\infty} \xi(A_n) |$$

(where $|\cdot|$ stands for the two-dimensional Euclidean norm), implies that the series $\Sigma_{n=1}^{\infty} \xi(A_n)$ is absolutely convergent.

The triple (Ω, Σ, ξ) is referred to as a complex measure space.

(ii) Given a complex measure space (Ω, Σ, ξ), the complex measure ξ is (explicitly) defined as $\xi = \nu_1 + i\nu_2$, where ν_1 and ν_2 are finite signed measures on Σ. Hahn decompositions should then be applied for ν_1 and ν_2 and their corresponding Jordan decompositions will yield

$$\xi = \nu_1^+ - \nu_1^- + i(\nu_2^+ - \nu_2^-), \tag{6.23}$$

with ν_1^+, ν_1^-, ν_2^+, and ν_2^- being positive finite measures. We will call ξ in (6.23) the *Jordan decomposition of the complex measure* ξ

and the positive measures ν_1^+, ν_1^-, ν_2^+, and ν^- the *Jordan variations*. □

7.2 Remark. It should be noted that the nature of a complex measure is different from that of a signed measure. The key reason is that the real and imaginary parts of a complex measure are represented by two arbitrary signed measures that carry out entirely different Hahn classes and Jordan variations. Bearing this in mind, we cannot expect the same characterizations between Jordan variations induced by Hahn decompositions and those in Proposition 6.21. For instance, if we define the total variation of ξ as $|\xi| = \nu_1^+ + \nu_1^- + \nu_2^+ + \nu_2^-$, we cannot expect it to agree with the characterization

$$|\xi|(A) = \sup\{\textstyle\sum_{k=1}^n |\xi(E_k)| : \{E_1, \ldots, E_n\} \subseteq \Sigma \text{ and } \sum_{k=1}^n E_k \subseteq A\} \tag{7.2}$$

holding for signed measures. Both are likely to be disconnected and consequently, we utilize the latter (rather than the former) to represent the total variation. So, we use (7.2) as the principal definition of the *total variation of a complex measure* ξ. □

Having a more consistent definition of the total variation for ξ we need to validate that the total variation is a measure, which we would not have to, should the representation of the total variation be in the form $|\xi| = \nu_1^+ + \nu_1^- + \nu_2^+ + \nu_2^-$.

We begin with three introductory statements very similar to Propositions 6.7 - 6 9 for signed measures. We thus leave their proofs to the reader (Problems 7.3 - 7.5).

7.3 Proposition. *Any complex measure ξ is continuous from below.*

7.4 Proposition. *Let (Ω, Σ, ξ) be a complex measure space and $\{A_n\}$ be a monotone nonincreasing sequence of measurable sets. Then the signed measure ξ is continuous from above on $\{A_n\}$.*

7.5 Proposition. *Let ξ be a finitely additive set function on the measurable space (Ω, Σ) such that $\xi(\emptyset) = 0$. ξ is continuous from below if and only if ξ is σ-additive.*

7.6 Proposition. *The total variation of a complex measure ξ on (Ω, Σ) is a finite positive measure on (Ω, Σ).*

Proof. First we notice that for an $E \in \Sigma$,

$$|\xi(E)| = \sqrt{[\nu_1(E)]^2 + [\nu_2(E)]^2} \leq |\nu_1(E)| + |\nu_2(E)|$$
$$\leq |\nu_1|(E) + |\nu_2|(E). \tag{7.6}$$

Here and throughout other numerous instances, the reader should be cautious not to confuse the Euclidean norm $|\cdot|$ with the total variation $|\cdot|$, because we use the same notation (running out of symbols). Now, let $A \in \Sigma$ and $E_1, \ldots, E_n$ be a measurable *subpartition* of A, meaning that $E_1, \ldots, E_n$ (given some $n \in \mathbb{N}$) are disjoint and $\sum_{i=1}^n E_i \subseteq A$. Then from (7.6),

$$\textstyle\sum_{i=1}^n |\xi(E_i)| \leq \sum_{i=1}^n |\nu_1|(E_i) + \sum_{i=1}^n |\nu_2|(E_i)$$

$$= |\nu_1|(\textstyle\sum_{i=1}^n E_i) + |\nu_2|(\sum_{i=1}^n E_i) \leq |\nu_1|(A) + |\nu_2|(A).$$

Obviously, the same inequality will hold for $\sup\{\sum_{k=1}^n |\xi(E_k)|\}$ of all subpartitions of A and thus

$$|\xi|(A) \leq |\nu_1|(A) + |\nu_2|(A). \tag{7.6a}$$

In conclusion, the latter inequality is the assurance that the set function $[\Sigma, \bar{\mathbb{R}}_+, |\xi|]$ defined so in (7.2) is reduced to $[\Sigma, \mathbb{R}_+, |\xi|]$, in as much as $|\nu_1|$ and $|\nu_2|$ are finite positive measures.

Obviously, $|\xi|(\emptyset) = 0$. Next we show that $|\xi|$ is a content on Σ (i.e., finitely additive). Let A and B be disjoint measurable sets and let $E_1, \ldots, E_n$ be a measurable subpartition of $A + B$. In as much as

$$\xi(E_i) = \xi(E_i \cap A) + \xi(E_i \cap A),$$

we have

$$|\xi(E_i)| \leq |\xi(E_i \cap A)| + |\xi(E_i \cap B)|$$

and thus

$$\sum_{i=1}^{n}|\xi(E_i)| \le \sum_{i=1}^{n}|\xi(E_i \cap A)| + \sum_{i=1}^{n}|\xi(E_i \cap B)|$$

(because $E_1 \cap A, \dots, E_n \cap A$ and $E_1 \cap B, \dots, E_n \cap B$ are measurable subpartitions of A and B)

$$\le |\xi|(A) + |\xi|(B).$$

The latter inequality holds true also for $\sup\{\sum_{k=1}^{n}|\xi(E_k)|\}$ of all measurable subpartitions of $A + B$. Therefore,

$$|\xi|(A + B) \le |\xi|(A) + |\xi|(B).$$

Now, we show the inverse inequality. Firstly we notice that if $E_1, \dots, E_n$ is a measurable subpartition of $A + B$,

$$F_1 : = E_1 \cap A, \dots, F_n : = E_n \cap A$$

and

$$F_{n+1} : = E_1 \cap B, \dots, F_{2n} : = E_n \cap B$$

are measurable subpartitions of A and B, respectively, and thus $F_1, \dots, F_{2n}$ is a subpartition of $A + B$. Hence,

$$\sum_{i=1}^{2n}|\xi(F_i)| \le |\xi|(A + B)$$

or in the form of

$$\sum_{i=1}^{n}|\xi(F_i)| + \sum_{i=n+1}^{2n}|\xi(F_i)| \le |\xi|(A + B)$$

implying that

$$|\xi|(A) + |\xi|(B) \le |\xi|(A + B)$$

and consequently, that $|\xi|$ is a content on Σ.

Finally, due to (7.3a), $|\xi|$ is continuous at the empty set and thus, its σ-additivity is due to Proposition 1.7(ii), Chapter 5. □

7.7 Remark. Although the set $\mathfrak{S}(\Omega, \Sigma)$ of all signed measures is not a vector space (the sum of two signed measures need not be a signed measure, as we can arrive at $\infty - \infty$), the space $\mathcal{C}(\Omega, \Sigma)$ of all complex measures (over field $\mathbb{C}$) is. To ensure that $\mathcal{C}(\Omega, \Sigma)$ is a vector space we proceed in the same way as for $\mathfrak{S}_*(\Omega, \Sigma)$, namely

that $\mathcal{C}(\Omega, \Sigma)$ is closed relative to linear combinations of complex measures with complex numbers. Next, the zero measure θ needs to be included in $\mathcal{C}(\Omega, \Sigma)$ as well. Furthermore, it can be shown that $\|\xi\|$ defined as $|\xi|(\Omega)$ is a norm on $\mathcal{C}(\Omega, \Sigma)$ (see Problem 7.10) and it therefore structures $\mathcal{C}(\Omega, \Sigma)$ as a normed vector space. Finally, Theorem 7.8 below (see Problem 7.13) claims that $(\mathcal{C}(\Omega, \Sigma), \|\cdot\|)$ is even a Banach space. □

7.8 Theorem. *The normed vector space* $(\mathcal{C}(\Omega, \Sigma), \|\cdot\|)$ *is Banach*. □

The following is an embellishment of the integral notion of real- and complex-valued functions with respect to signed and complex measures.

7.9 Definitions.

(i) **The integral of a complex-valued function with respect to a positive measure**. Let $[\Omega, \mathbb{C}, f = u + iv]$ be a complex-valued function. Given a σ-algebra Σ in Ω, f is *measurable*, if for every Borel set $B \in \mathcal{B}(\mathbb{R}^2)$, $f^*(B) \in \Sigma$, as usual. By using the projection operators one can show that f is Σ-$\mathcal{B}(\mathbb{R}^2)$ measurable if and only if u and v are Σ-$\mathcal{B}(\mathbb{R})$ measurable. (More about it in Section 9 on product measures.) Now, given a positive measure $\mu \in \mathfrak{M}(\Omega, \Sigma)$, we say that $f \in L^1(\Omega, \Sigma, \mu; \mathbb{C})$ or f is *μ-integrable* if $|f| \in L^1(\Omega, \Sigma, \mu; \mathbb{R}_+)$.

As $|f| \leq |u| + |v| \leq 2|f|$, f is integrable if and only if both u and v are elements of $L^1(\Omega, \Sigma, \mu; \mathbb{R})$ and in this case we will write

$$\int f d\mu = \int u d\mu + i \int v d\mu. \tag{7.9}$$

Therefore, $L^1(\Omega, \Sigma, \mu; \mathbb{C})$ is a linear space with the integral being a linear functional on $L^1(\Omega, \Sigma, \mu; \mathbb{C})$. All major theorems of integration (cf. Sections 1 and 2) hold true with very minor notational modifications.

(ii) **The integral of an extended real-valued function with respect to a signed measure**. Let $\nu \in \mathfrak{S}(\Omega, \Sigma)$ with its Jordan decomposition $\nu = \nu^+ - \nu^-$. Denote

$$L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}}) = L^1(\Omega, \Sigma, \nu^+; \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma, \nu^-; \overline{\mathbb{R}}). \tag{7.9a}$$

The *integral of a function* $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ *relative to the signed measure* ν is then defined as

$$\int f d\nu = \int f d\nu^+ - \int f d\nu^-. \tag{7.9b}$$

The function f is said to be *integrable with respect to the signed measure* ν, if $f \in L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}})$. The value of the integral $\int f d\nu$ is then finite. (For further discussion, see Problem 7.15.)

Notice that the full decomposition of (7.9b) is

$$\int f d\nu = \int f^+ d\nu^+ + \int f^- d\nu^- - \int f^- d\nu^+ - \int f^+ d\nu^- \tag{7.9c}$$

and thus the space $L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}})$ could be enlarged by including those f's whose one of the two parts, "positive" $\int f^+ d\nu^+ + \int f^- d\nu^-$ or "negative" $\int f^- d\nu^+ + \int f^+ d\nu^-$ in (7.9c) is finite. In this case, we say that the integral $\int f d\nu$ of f exists and we write $f \in \mathbb{L}(\Omega, \Sigma, \nu; \overline{\mathbb{R}})$.

(iii) **The integral of an extended real-valued function with respect to a complex measure**. Let $\xi = \nu_1 + i\nu_2$ be a complex measure on a measurable space (Ω, Σ). Denote

$$L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}) = L^1(\Omega, \Sigma; \nu_1, \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma; \nu_2, \overline{\mathbb{R}}) \tag{7.9d}$$

bearing in mind (7.9a). Then, define the integral of an extended real-valued integrable function $f \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}})$ as

$$\int f d\xi = \int f d\nu_1 + i \int f d\nu_2. \tag{7.9e}$$

Unlike the previous case (ii), we do not attempt to enlarge the class of integrable functions for the simple reason that $\int f d\xi$ must also represent a complex measure and thus it must be finite.

(iv) **The integral of a complex-valued function with respect to a complex measure**. Let $\xi = \nu_1 + i\nu_2 \in \mathcal{C}(\Omega, \Sigma)$ and let $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{C})$ such that $f = u + iv$ and $u, v \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}})$ of (7.9d). Then, we say that f is *integrable with respect to the complex measure* ξ and write $f \in L^1(\Omega, \Sigma; \xi, \mathbb{C})$. Then, the integral of f is defined as

$$\int f d\xi = \int u d\nu_1 - \int v d\nu_2 + i(\int v d\nu_1 + \int u d\nu_2), \tag{7.9f}$$

□

7.10 Example. Replacing f with $f\mathbf{1}_A$ we see that (7.9f) also defines a complex measure and therefore $\xi \mapsto \int f d\xi$ is a linear operator from $\mathcal{C}(\Omega, \Sigma)$ to $\mathcal{C}(\Omega, \Sigma)$.

Now, let $\mathcal{C}_*^{-1}(\Omega, \Sigma; \mathbb{C})$ denote the linear space of all Σ-measurable complex-valued bounded functions. Let $\xi = \nu_1 + i\nu_2$ be a complex measure on (Ω, Σ). The integral in (7.9f), as a functional of f, is clearly linear.

We define a norm on $\mathcal{C}_*^{-1}(\Omega, \Sigma; \mathbb{C})$ as

$$\|f\|_\infty = \sup\{|f(\omega)| : \omega \in \Omega\}.$$

Given a $\xi \in \mathcal{C}(\Omega, \Sigma)$, and s being a simple function,

$$s = \textstyle\sum_{k=1}^n a_k \mathbf{1}_{A_k},$$

we have

$$\begin{aligned} |\int s d\xi| &= |\textstyle\sum_{k=1}^n a_k \xi(A_k)| \\ &\leq \textstyle\sum_{k=1}^n a_k |\xi(A_k)| \leq \sum_{k=1}^n \|s\|_\infty |\xi(A_k)| \\ &\leq \|s\|_\infty \, \|\xi\|. \end{aligned}$$

If $f \in \mathcal{C}_*^{-1}(\Omega, \Sigma; \mathbb{R}_+)$, then there is a sequence $\{s_n\}$ of simple functions with $s_n \uparrow f$. Hence,

$$|\int s_n d\xi| \leq \|s_n\|_\infty \, \|\xi\| \leq \|f\|_\infty \, \|\xi\|,$$

which leads to

$$|\int f d\xi| \leq \|f\|_\infty \|\xi\| < \infty$$

and obviously, to the same inequality for $f \in \mathcal{C}_*^{-1}(\Omega, \Sigma; \mathbb{C})$. The latter implies that $\mathcal{C}_*^{-1}(\Omega, \Sigma; \mathbb{C}) \subseteq L^1(\Omega, \Sigma; \xi, \mathbb{C})$. □

PROBLEMS

7.1 Show that the total variation of a complex measure ξ has the property $|\xi| \leq \nu_1^+ + \nu_1^- + \nu_2^+ + \nu^- = |\nu_1| + |\nu_2|$, and that in general, $|\xi| < \nu_1^+ + \nu_1^- + \nu_2^+ + \nu^-$. Thus, it disagrees with the associated property for a signed measure.

7.2 Prove the statement: *Let (Ω, Σ, ξ) be a complex measure space. If A and B are measurable sets such that $A \subseteq B$, then $\xi(B \backslash A) = \xi(B) - \xi(A)$.*

7.3 Prove Proposition 7.3. *Any complex measure ξ is continuous from below.*

7.4 Prove Proposition 7.4. *Let (Ω, Σ, ξ) be a complex measure space and $\{A_n\}$ be a monotone nonincreasing sequence of measurable sets. Then the signed measure ξ is continuous from above on $\{A_n\}$.*

7.5 Prove Proposition 7.5. *Let ξ be a finitely additive set function on the measurable space (Ω, Σ) such that $\xi(\emptyset) = 0$. ξ is continuous from below if and only if ξ is σ-additive.*

7.6 Show that for any complex measure ξ, $|\xi(A)| \leq |\xi|(A)$, for all $A \in \Sigma$.

7.7 Any complex measure $\xi \in \mathcal{C}(\Omega, \Sigma)$ is finite. It means that for any measurable set A, $|\xi(A)| < \infty$. (a) Does it mean that ξ is

bounded? (b) Let $\xi = \nu_1 + i\nu_2$, where $\nu_1, \nu_2 \in \mathfrak{S}_*(\Omega, \Sigma)$. Show that for any $A \in \Sigma$, $|\xi(A)| \leq \|\nu_1\| + \|\nu_2\|$.

7.8 Prove that for any complex measure $\xi = \nu_1 + i\nu_2$, $\|\xi\| \leq \|\nu_1\| + \|\nu_2\|$.

7.9 Let $\xi, \xi_1, \xi_2 \in \mathcal{C}(\Omega, \Sigma)$ such that for each $A \in \Sigma$, $|\xi(A)| \leq |\xi_1(A)| + |\xi_2(A)|$. Prove that $\|\xi\| \leq \|\xi_1\| + \|\xi_2\|$.

7.10 Show that the real-valued function $\|\cdot\|$ introduced in Remark 7.7 is a norm on $\mathcal{C}(\Omega, \Sigma)$. Namely, let (Ω, Σ, ξ) be a complex measure space and let $\|\xi\| := |\xi|(\Omega)$ be a real-valued function defined on $\mathcal{C}(\Omega, \Sigma)$. Then, $\|\cdot\|$ is a norm.

7.11 Let $\xi = \nu + i\rho$ be a complex measure on a measurable space (Ω, Σ). Prove the following inequalities for total variations and norms.

$$|\xi| \geq \tfrac{1}{\sqrt{2}}(|\nu| + |\rho|) \text{ and } \|\xi\| \geq \tfrac{1}{\sqrt{2}}(\|\nu\| + \|\rho\|).$$

7.12 Let $\{\xi_n = \nu_n + i\rho_n\}$ be a sequence of complex measures on a measurable space (Ω, Σ). Show that $\{\xi_n\}$ is a Cauchy sequence if and only if $\{\nu_n\}$ and $\{\rho_n\}$ are Cauchy sequences.

7.13 Prove Theorem 7.8 that $(\mathcal{C}(\Omega, \Sigma), \|\cdot\|)$ is a Banach space.

7.14 Show that for each $f \in L^1(\Omega, \Sigma, \mu; \mathbb{C})$, $|\int f d\mu| \leq \int |f| d\mu$, where μ is a positive measure.

7.15 Show that, given a signed measure ν, it holds true that $L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}}) = L^1(\Omega, \Sigma, |\nu|; \overline{\mathbb{R}})$.

7.16 Modify the Lebesgue dominated convergence theorem of Section 2, Chapter 6, and prove it for complex-valued functions.

7.17 Let (Ω, Σ, μ) be a measure space, with μ being a positive measure. Prove that for each $f \in L^1(\Omega, \Sigma, \mu; \mathbb{C})$, $|\xi| = \int |f| d\mu$, where $\xi = \int f d\mu$.

7.18 Let (Ω, Σ, μ) be a positive measure space and let $f \in \mathcal{C}^{-1}(\mathbb{C})$. Show that $\int\limits_A f d\mu = (0,0)$ for each $A \in \Sigma$ if and only if $f = (0,0)$ μ-a.e. on Ω.

7.19 Let (Ω, Σ, ν) be a signed measure space and $f \in \mathbb{L}(\Omega, \Sigma; \nu, \overline{\mathbb{R}}_+)$. Define $\rho = \int f d\nu^+ - \int f d\nu^-$, $\rho_1 = \int f d\nu^+$, and $\rho_2 = \int f d\nu^-$. Prove that ρ_1 and ρ_2 are the Jordan variations of ρ.

7.20 Let (Ω, Σ, ν) be a signed measure space and $f \in \mathbb{L}(\Omega, \Sigma; \nu, \overline{\mathbb{R}}_+)$. Prove that $f = 0$ ν-a.e. if and only if $\int\limits_A f d\nu = 0$ for all $A \in \Sigma$.

7.21 Let (Ω, Σ, ν) be a signed measure space and $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$. Prove that $f = 0$ ν-a.e. if and only if $\int\limits_A f d\nu = 0$ for all $A \in \Sigma$.

7.22 Let (Ω, Σ, ν) be a signed measure space and $r \in \mathbb{R}$ (a real constant). Show that $r\nu$ is a signed measure on Σ and prove that $L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}}) = L^1(\Omega, \Sigma; r\nu, \overline{\mathbb{R}})$ and that $\int f d(r\nu) = r \int f d\nu$.

7.23 Prove the monotone convergence theorem for a monotone nondecreasing sequence of nonnegative measurable functions $\{f_n\}$ on a signed measure space pointwise convergent to $f \in \mathbb{L}(\Omega, \Sigma; \nu, \overline{\mathbb{R}}_+)$, namely,

$$\lim_{n\to\infty} \int f_n d\nu = \int f d\nu.$$

7.24 Let (Ω, Σ) be a measurable space and ν and ρ be two signed measures on (Ω, Σ). Suppose $\nu + \rho$ is defined. Show that it is

a signed measure on (Ω, Σ) and prove that for any $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma; \rho, \overline{\mathbb{R}})$,

$$\int f d(\nu + \rho) = \int f d\nu + \int f d\rho.$$

7.25 Let (Ω, Σ) be a measurable space and ν and ρ be two signed measures on (Ω, Σ) such that $\nu + \rho$ is defined. Suppose a and b are two real constants. (i) Is the linear combination $a\nu + b\rho$ always defined? (ii) If $a\nu + b\rho$ is defined, show that it is a signed measure on (Ω, Σ) and prove that for any $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma; \rho, \overline{\mathbb{R}})$,

$$\int f d(a\nu + b\rho) = a \int f d\nu + b \int f d\rho.$$

7.26 Let (Ω, Σ, ξ) be a complex measure space and $c \in \mathbb{C}$ (a complex number). Show that $c\xi$ is a complex measure on (Ω, Σ) and prove that $L^1(\Omega, \Sigma; \xi, \mathbb{C}) = L^1(\Omega, \Sigma; c\xi, \mathbb{C})$ and that $\int f d(c\xi) = c \int f d\xi$.

7.27 Let (Ω, Σ) be a measurable space and ξ and η be two complex measures on (Ω, Σ). Show that $\xi + \eta$ is a complex measure on (Ω, Σ) and prove that for any $f \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma; \eta, \overline{\mathbb{R}})$,

$$\int f d(\xi + \eta) = \int f d\xi + \int f d\eta.$$

7.28 Modify Problem 7.26 under the assumption that $f \in L^1(\Omega, \Sigma; \xi, \mathbb{C}) \cap L^1(\Omega, \Sigma; \eta, \mathbb{C})$.

7.29 Let $\xi, \eta \in \mathcal{C}(\Omega, \Sigma)$ and α, β be two complex numbers. Prove that for any $f \in L^1(\Omega, \Sigma; \xi, \mathbb{C}) \cap L^1(\Omega, \Sigma; \eta, \mathbb{C})$,

$$\int f d(\alpha\xi + \beta\eta) = \alpha \int f d\xi + \beta \int f d\eta.$$

7.30 Let (Ω, Σ, ν) be a signed measure space and $f \in L^1(\Omega, \Sigma; \nu, \mathbb{C})$. Prove that $f = 0$ ν-a.e. if and only if $\int\limits_A f d\nu = 0$ for all $A \in \Sigma$.

7.31 Let (Ω, Σ, ξ) be a complex measure space and $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$. Prove that $f = 0$ ξ-a.e. if and only if $\int_A f d\xi = 0$ for all $A \in \Sigma$.

7.32 Let $\nu \in \mathfrak{S}_*(\Omega, \Sigma)$. Show that $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$ if and only if $|f| \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}}_+)$.

7.33 Show that $(L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}}), +, \cdot, \preceq)$ is a vector lattice over $\mathbb{R}$.

7.34 Let $\xi \in \mathcal{C}(\Omega, \Sigma)$ (be a complex measure). Show that a function $f \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}})$ if and only if $|f| \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}_+)$. Also, define $L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}_+)$.

7.35 Let $\xi \in \mathcal{C}(\Omega, \Sigma)$ (be a complex measure). Show that a function $f \in L^1(\Omega, \Sigma; \xi, \mathbb{C})$ if and only if $|f| \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}_+)$.

NEW TERMS:

complex measure, 458
complex measure space, 458
Jordan decomposition of a complex measure, 458
total variation of a complex measure, 459
measurable subpartition of a set, 460
$\mathcal{C}(\Omega, \Sigma)$, vector space of all complex measures over field $\mathbb{C}$, 462
$\|\xi\|$, norm of a complex measure $\mathcal{C}(\Omega, \Sigma)$, 462
$(\mathcal{C}(\Omega, \Sigma), \|\cdot\|)$, as a Banach space, 462
measurability of a complex-valued function, 462
integral of a complex-valued functions with respect to a positive measure, 462
integrability of a complex-valued function, 462
integral of an extended real-valued function with respect to a signed measure, 463
space $\mathbb{L}(\Omega, \Sigma, \nu; \overline{\mathbb{R}})$, 463
space $L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}})$, 463
integral of an extended real-valued function with respect to a complex measure, 463
integral of a complex-valued function with respect to a complex measure, 464
Lebesgue dominated convergence theorem for complex-valued functions, 466
monotone convergence theorem on a signed measure space, 467

8. RADON - NIKODYM THEOREM FOR SIGNED AND COMPLEX MEASURES AND ITS APPLICATIONS

We continue our discussion on absolute continuity started in Section 5 for positive measures. In this section we present a more rigorous version of the Radon - Nikodym theorem, along with detailed proofs and some common applications to probability theory, namely, the conditional expectation.

8.1 Definition and Notation. Absolute continuity of signed measures is formulated in the same way as that of positive measures. Let μ and ξ be positive and signed or complex measures, respectively, on a measure space (Ω, Σ). We call measure ξ *absolutely continuous with respect to measure* μ (in notation, $\nu \ll \mu$) if every measurable μ-null set A is also a ξ-null set. A signed or complex measure that is absolutely continuous with respect to the Lebesgue measure is called just *continuous*.

A signed or complex measure ξ is said to be absolutely continuous with respect to another signed or complex measure η, in notation $\xi \ll \eta$, if $\xi \ll |\eta|$. □

We start this section with some easy-to-prove preliminaries.

8.2 Proposition. *Let ν be a signed measure and μ be a positive measure on a measurable space (Ω, Σ). Then the following statements hold*:

(i) $\nu \ll \mu \Rightarrow \nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow |\nu| \ll \mu$.

(ii) $\nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow \nu \ll \mu$ and $|\nu| \ll \mu$.

(iii) $|\nu| \ll \mu \Rightarrow \nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow \nu \ll \mu$.

8.3 Proposition. *Let $\xi = \nu + i\rho$ be a complex measure and μ be a positive measure on a measurable space (Ω, Σ). Then, the following statements are equivalent*.

(i) $\xi \ll \mu$.

(ii) $|\xi| \ll \mu$.

(iii) $\nu^+ \ll \mu,\ \nu^- \ll \mu,\ \rho^+ \ll \mu,\ \rho^- \ll \mu,\ \nu \ll \mu,\ \rho \ll \mu.$

(See Problems 8.1 and 8.2.) □

8.4 Notation. Recall that $\mathfrak{S}(\Omega, \Sigma)$ denotes the set of all signed measures on (Ω, Σ) and $\mathfrak{M}(\Omega, \Sigma)$ is the subset of all positive measures. Given $\mu \in \mathfrak{M}$, denote $\mathfrak{S}_\mu^{\ll} = \{\nu \in \mathfrak{S}(\Omega, \Sigma) : \nu \ll \mu\}$. Define on the space $\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ the map $\int_\mu$ such that for each $g \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$,

$$\int_\mu g \,(= \int g d\mu) \ = \nu \tag{8.4}$$

that, according to Problem 6.1, is valued in $\mathfrak{S}$ and, as easily seen, $\nu \ll \mu$. Consequently, $[\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})\,, \mathfrak{S}_\mu^{\ll}, \int_\mu]$ is an into map. □

From Definitions and Remarks 5.7(iii), Chapter 5, we recollect that the μ-almost everywhere property of equality of measurable functions generates the equivalence kernel modulo E_μ (or just *modulo* μ) on $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ and thus on its subset $\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. Consequently,

$$\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_{\mathrm{E}_\mu} \ = \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu$$

is also a quotient set. On the other hand, by Corollary 1.20, Chapter 6, the map $\int_\mu$ "agrees" with E_μ. In other words, $\int_\mu$ adopts E_μ as its equivalence kernel. Then, by Theorem 4.4, Chapter 1, there is a unique reducer of $\int_\mu$, denoted by

$$[\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu, \mathfrak{S}_\mu^{\ll}, \boldsymbol{\int_\mu}],$$

such that

$$\int_\mu = \boldsymbol{\int_\mu} \circ \pi_{\mathrm{E}},$$

where π_{E} is the projection of $\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ on its quotient $\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu$ by E_μ. Therefore, $\int_\mu$ turns to an injective map $\boldsymbol{\int_\mu}$ with the domain $\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu$.

Recall that any function g from the quotient class $[g]_\mu$ of $\mathbb{L}$-functions [generating the signed measure ν in (8.1)] is referred to as a Radon - Nikodym *density of* (*the signed measure*) ν *relative to* (*the positive measure*) μ.

As in Section 5 for positive measures, the chief factor of the Radon - Nikodym theorem below is that, under some minor restriction imposed on μ, the mapping $[\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu, \mathfrak{S}_\mu^{\ll}, \int_{\boldsymbol{\mu}}]$ is surjective. In other words, for each measure $\nu \in \mathfrak{S}_\mu^{\ll}$, there is an equivalence class $[g]_\mu$ of Radon - Nikodym densities of ν relative to μ.

8.5 Theorem (Radon - Nikodym)**.** *Let $\mu \in \mathfrak{M}(\Omega, \Sigma)$ be a σ-finite measure. Then $[\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu, \mathfrak{S}_\mu^{\ll}, \int_{\boldsymbol{\mu}}]$ is a bijective map.*

Proof. The proof of the theorem includes three objectives.

1) Show that given $[g]_\mu \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu$, $\int_{\boldsymbol{\mu}}[g]_\mu \in \mathfrak{S}_\mu^{\ll}$, that is, for each $g \in [g]_\mu$, $\nu = \int g d\mu$ defines a signed measure, absolutely continuous with respect to μ. This is done. (Because $g \in \mathbb{L}$, ν is a signed measure as per Problem 6.1. The proof that $\nu \ll \mu$ is trivial.) Therefore,

$$\int_{\boldsymbol{\mu}}: \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu \to \mathfrak{S}_\mu^{\ll}$$

is an into map.

2) Show that

$$\int_{\boldsymbol{\mu}}: \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu \to \mathfrak{S}_\mu^{\ll}$$

is a surjective (an onto) map, that is, for every signed measure $\nu \in \mathfrak{S}_\mu^{\ll}$, which is absolutely continuous with respect to a positive σ-finite measure μ, there is an equivalent class $[g]_\mu \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu$ of Radon - Nikodym densities of ν relative to μ. This is the key part of the theorem and it is referred to as the "existence of a Radon - Nikodym density."

3) Show that the map

$$\int_{\mu}: \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_{\mu} \ \rightarrow \ \mathfrak{S}_{\mu}^{\ll}$$

is injective (one-to-one); that is, the above equivalent class $[g]_{\mu}$ is unique. This is due to Corollary 1.20.

It therefore remains to prove the existence of the Radon-Nikodym density, namely, given a signed measure ν, absolutely continuous relative to a σ-finite positive measure μ, there is an equivalence class $[g]_{\mu}$ of $\mathbb{L}$-functions such that for each element g of this class $\nu = \int g d\mu$. We break up the proof of the existence into five parts starting with the case of two finite measures μ and ν and embellishing it to the case when μ and ν are σ-finite positive and signed measures, respectively.

Case 1: μ and ν are two finite positive measures

Abbreviate $L_{+}^{1} = L^{1}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_{+})$ and introduce the subset of functions

$$\Phi = \left\{ f \in L_{+}^{1}: \int_{A} f d\mu \leq \nu(A), \forall A \in \Sigma \right\}. \tag{8.5}$$

Because $\mathbf{0} \in \Phi$, the latter is not empty. Furthermore, Φ is closed under finite suprema. Indeed, let $f, g \in \Phi$ and $A \in \Sigma$. Denote by

$$E = \{\omega \in A: f(\omega) \geq g(\omega)\}$$

and

$$G = \{\omega \in A: f(\omega) < g(\omega)\}.$$

Then, $E + G = A$ and

$$\int_{A} f \vee g d\mu = \int_{E} f d\mu + \int_{G} g d\mu \leq \nu(E) + \nu(G) = \nu(A).$$

Now, let

$$S: = \sup \left\{ \int_{\Omega} f d\mu: f \in \Phi \right\} \leq \nu(\Omega) < \infty.$$

Then, there is a sequence $\{\varphi_n\} \subseteq \Phi$ such that $\sup\{\int \varphi_n \, d\mu\} = S$. Indeed, because $S < \infty$, for each $n = 1, 2, \ldots,$ there is a function φ_n such that $S - \frac{1}{n} \leq \int \varphi_n d\mu \leq S$. By setting $f_n = \bigvee_{i=1}^{n} \varphi_i$ we form the new sequence $\{f_n\}$, which is monotone nondecreasing and has S as $\lim_{n\to\infty} \int f_n d\mu$. By the monotone convergence theorem we have that $g := \lim_{n\to\infty} f_n$, in the topology of pointwise convergence, and is an L^1_+-function that obviously belongs to Φ. The function g we just arrived at is an $\int$-maximal element of Φ and it is an element of the equivalent class $[g]_\mu$ of $\int$-maximal elements. (For a pertinent discussion see Problem 8.11.)

Now, we show that $[g]_\mu$ is an equivalence class of Radon - Nikodym densities of ν relative to μ, that is, for each $g \in [g]_\mu$,

$$\int_A g d\mu = \nu(A), \quad \text{for all } A \in \Sigma.$$

Because $\nu \ll \mu$ and because for all $A \in \Sigma$ (from the definition of Φ)

$$\int_A g d\mu \leq \nu(A),$$

the set function

$$\rho = \nu - \int g d\mu$$

is a finite positive measure, absolutely continuous relative to measure μ. If g is not a Radon - Nikodym density of ν with respect to μ, then $\rho \not\equiv 0$ and thus $\rho(\Omega) > 0$. The latter implies that, for some positive ε,

$$\mu(\Omega) - \varepsilon\rho(\Omega) < 0. \tag{8.5a}$$

Consider the (finite) signed measure $\gamma = \mu - \varepsilon\rho$. By Theorem 6.12, there is a Hahn decomposition (P, N) of Ω such that $\gamma(A \cap P) \geq 0$ and $\gamma(A \cap N) \leq 0$ for all measurable sets A. In other words,

$$\mu(P \cap A) - \varepsilon\rho(P \cap A) \geq 0 \tag{8.5b}$$

and

$$\mu(N \cap A) - \varepsilon\rho(N \cap A) \leq 0. \tag{8.5c}$$

If $\mu(N) = 0$ then, because of $\rho \ll \mu$, $\rho(N) = 0$ and thus $\gamma(N) = 0$. On the other hand, from (8.5b), by setting $A = \Omega$ we have that

$$\mu(P) - \varepsilon\rho(P) = \gamma(P) \geq 0\,. \tag{8.5d}$$

Furthermore, because by the above assumption about ρ, N turns out to be a γ-null set, it follows from (8.5a) that $\gamma(P) < 0$. This contradicts inequality (8.5d). Hence, $\mu(N)$ must be positive.

Now we have from (8.5c) that

$$\int_A \tfrac{1}{\varepsilon}\,\mathbf{1}_N d\mu = \tfrac{1}{\varepsilon}\,\mu(A \cap N)$$

$$\leq \rho(N \cap A) \leq \rho(A) = \nu(A) - \int_A g d\mu$$

or, equivalently,

$$\int_A \big(\tfrac{1}{\varepsilon}\,\mathbf{1}_N + g\big) d\mu \leq \nu(A)\,.$$

Thus, the function $\frac{1}{\varepsilon}\,\mathbf{1}_N + g \in \Phi$. But, because $\mu(N) > 0$, it holds true that

$$\int \big(\tfrac{1}{\varepsilon}\mathbf{1}_N + g\big) d\mu > \int g d\mu.$$

This contradicts that g is an $\int$-maximal element of Φ. Evidently, the contradiction is due to the wrong assumption about ρ. Thus $\rho \equiv 0$, or, in other words,

$$\nu(A) = \int_A g d\mu$$

for all $A \in \Sigma$, which proves the statement of the theorem for this special case.

Notice that because ν is finite and thus every Radon - Nikodym density g is an L^1-function, by Proposition 1.21, Chapter 6, g is finite μ-a.e. If it is "occasionally" infinite, we can redefine g to make it finite. Therefore, in the equivalence class $[g]_\mu$ of Radon - Nikodym densities there is a subclass of finite densities.

To summarize case 1, given two finite positive measures μ and ν such that $\nu \ll \mu$, there is a unique (nonempty) equivalence class $[g]_\mu \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+)|_\mu$ of Radon - Nikodym densities (of measure ν relative to measure μ), with a nonempty subclass of finite densities.

Case 2. μ and ν are finite and σ-finite positive measures, respectively

If ν is σ-finite then there is at most a countable decomposition of $\Omega = \Sigma_{n=1}^{\infty}\Omega_n$, such that $\nu(\Omega_n) < \infty$ for all $n = 1, 2, \dots$. Let

$$\nu_n = \mathrm{Res}_{\Sigma\cap\Omega_n}\nu.$$

Then ν_n is a finite measure on $\Omega_n \cap \Sigma$ and from case 1 it follows that there is a measurable nonnegative function $g_n\colon \Omega_n \to \mathbb{R}$ such that

$$\nu_n(A \cap \Omega_n) = \int\limits_{A\cap\Omega_n} g_n \, d\mu\,, \quad \text{for each } A \in \Sigma,\ n = 1, 2, \dots\ .$$

Now by the monotone convergence theorem applied to the sequence $\{\Sigma_{n=1}^{m} g_n \mathbf{1}_{\Omega_n}\}$ we have that

$$\begin{aligned}\nu(A) &= \textstyle\sum_{n=1}^{\infty} \nu_n(A \cap \Omega_n) \\ &= \textstyle\sum_{n=1}^{\infty} \int\limits_{A\cap\Omega_n} g_n d\mu = \int\limits_A \sum_{n=1}^{\infty} g_n \mathbf{1}_{\Omega_n} d\mu\,.\end{aligned}$$

It only remains to set $g = \Sigma_{n=1}^{\infty} g_n \mathbf{1}_{\Omega_n}$ to complete this part of the theorem.

Therefore, given two positive measures μ and ν such that μ is finite, ν is σ-finite and $\nu \ll \mu$, there is a unique equivalence class $[g]_\mu \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+) \mid_\mu$ of Radon - Nikodym densities (whose integral is not necessarily finite).

Case 3. μ is a finite positive and ν is an arbitrary positive measure

Denote by

$$\Gamma = \{B \in \Sigma\colon \mathrm{Res}_{\Sigma\cap B}\nu \text{ is } \sigma\text{-finite}\}.$$

In as much as $\emptyset \in \Gamma$, it follows that $\Gamma \neq \emptyset$. Let

$$S = \sup\{\mu(B)\colon B \in \Gamma\}$$

and let $\{E_n\} \subseteq \Gamma$ such that $\mu(E_n) \to S$. ($S \leq \mu(\Omega) < \infty$, for each $n = 1, 2, \ldots$. Therefore, there is a set E_n such that $S - \frac{1}{n} \leq \mu(E_n) \leq S$.) Clearly,

$$E\colon = \bigcup_{n=1}^{\infty} E_n \in \Gamma.$$

Hence, $S \geq \mu(E) \geq \mu(E_n) \to S$ and $\mu(E) = S$.

Now because ν is σ-finite on E, from case 2, it follows that there is a $\Sigma \cap E$-$\overline{\mathcal{B}}_+$-measurable L^1-function $[E, \overline{\mathbb{R}}_+, \widetilde{g}]$, such that

$$\nu_E(A) = \nu(A \cap E) = \int_{A \cap E} \widetilde{g}\, d\mu$$

for all $A \in \Sigma$. Fix an $A \in \Sigma$.

a) Let $\mu(A \cap E^c) > 0$. If $\nu(A \cap E^c) < \infty$, then $A \cap E^c \in \Gamma$, and thus $E \cup (A \cap E^c) \in \Gamma$. The latter yields that

$$S \geq \mu(E + (A \cap E^c)) = \mu(E) + \mu(A \cap E^c) > \mu(E)$$

and this contradicts $\mu(E) = S$. Thus $\nu(A \cap E^c) = \infty$.

b) Let $\mu(A \cap E^c) = 0$. Then because $\nu \ll \mu$, it holds true that $\nu(A \cap E^c) = 0$.

The above cases *a*) and *b*) can be combined in one compact equation

$$\nu(A \cap E^c) = \int \infty\, 1_{A \cap E^c}\, d\mu$$

by agreeing that $\infty \cdot 0 = 0$. Furthermore,

$$\nu(A) = \nu(A \cap E) + \nu(A \cap E^c)$$

$$= \int_{A \cap E} \widetilde{g}\, d\mu + \int \infty\, \mathbf{1}_{A \cap E^c}\, d\mu = \int_A g d\mu\,,$$

where $g = \widetilde{g}\ \mathbf{1}_E + \infty \mathbf{1}_{E^c}$. Notice that g is measurable, because

$$\sigma((\Sigma \cap E) \cup (\Sigma \cap E^c)) \subseteq \Sigma.$$

Therefore, given two positive measures μ and ν such that μ is finite, ν is arbitrary, and $\nu \ll \mu$, there is unique equivalence class $[g]_\mu \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+) \mid_\mu$ of Radon - Nikodym densities.

Case 4. μ is a σ-finite and ν is an arbitrary positive measure.

Let $\Omega = \Sigma_{n=1}^{\infty} \Omega_n$ such that $\mu(\Omega_n) < \infty$ for all $n \geq 1$. Due to Case 3, for each n, there is a $\Sigma \cap \Omega_n$-$\overline{\mathcal{B}}_+$-measurable function $[\Omega_n, \overline{\mathbb{R}}_+, \widetilde{g}_n]$, such that

$$\nu(A \cap \Omega_n) = \int_{A \cap \Omega_n} \widetilde{g}_n \, d\mu,$$

for all $A \in \Sigma$. Denoting $g_n = \widetilde{g}\, \mathbf{1}_{\Omega_n}$ we have

$$\nu(A \cap \Omega_n) = \int\limits_A g_n \, d\mu$$

and thus

$$\nu(A) = \int\limits_A g d\mu,$$

where, by the monotone convergence theorem, $g = \Sigma_{n=1}^{\infty} g_n$.

Therefore, given two positive measures μ and ν of which μ is σ-finite and ν is arbitrary and $\nu \ll \mu$, there is a unique equivalence class $[g]_\mu \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+)|_\mu$ of (nonnegative) Radon - Nikodym densities.

Case 5. μ is a σ-finite positive measure and ν is a signed measure

Let $\nu = \nu^+ - \nu^-$ be the Jordan decomposition of ν, where, for instance, ν^- is supposed to be finite. By Case 4, there are functions $[\Omega, \overline{\mathbb{R}}_+, g_i]$ each belonging to their respective equivalence classes $[g_i]_\mu$, $i = 1, 2$, such that

$$\nu^+(A) = \int\limits_A g_1 d\mu \ \text{ and } \ \nu^-(A) = \int\limits_A g_2 \, d\mu, \ A \in \Sigma.$$

In as much as by our assumption ν^- is finite, g_2 is μ-integrable and $g_2 < \infty$ μ-a.e. and $g := g_1 - g_2$ makes sense. This leads to

$$\nu(A) = \nu^+(A) - \nu^-(A) = \int_A g d\mu.$$

Thus, we can say that g belongs to an equivalence class

$$[g]_\mu = [g_1]_\mu - [g_2]_\mu = \left\{ g_1 - g_2 : g_1 \in [g_1]_\mu, g_2 \in [g_2]_\mu \right\}$$

of Radon - Nikodym densities.

In summary of Case 5, given a σ-finite positive measure μ and signed measure ν, with $\nu \ll \mu$, there is a unique equivalence class $[g]_\mu \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu$ of Radon - Nikodym densities of ν relative to μ.

Case 5a (special case of 5; ν is a finite signed measure.)

In this case, clearly, given a σ-finite positive measure μ and a finite signed measure ν, with $\nu \ll \mu$, there is a unique equivalence class $[g]_\mu \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}) \mid_\mu$ of Radon - Nikodym densities of ν relative to μ.

Case 5b (special case of 5; ν is a σ-finite signed measure.)

In as much as ν is σ-finite, there is a countable decomposition $\Sigma_{n=1}^\infty \Omega_n = \Omega$ such that $\nu_n = \mathrm{Res}_{\Sigma \cap \Omega_n} \nu$ is finite for every n. By Case 5a, because

$$\nu_n \ll \mu_n = \mathrm{Res}_{\Sigma \cap \Omega_n} \mu,$$

there is a unique equivalence class $[g_n]_\mu \in L^1(\Omega_n, \Sigma \cap \Omega_n, \mu_n; \overline{\mathbb{R}}) \mid_\mu$ of Radon - Nikodym densities of ν_n relative to μ_n. Now, letting

$$g = \sum_{n=1}^\infty 1_{\Omega_n} g_n,$$

for every $g_n \in [g_n]_{\mu_n}$, we define the class $[g]_\mu$ of Radon - Nikodym densities of ν relative to μ. As a sum of countably many integrable functions, g is clearly μ-a.e. finite.

The proof of the theorem is now complete. □

8.6 Definition and Notation. By the Radon - Nikodym theorem, the map $\int_{\mu}$ is therefore invertible and its inverse $\int_{\mu}^{-1}$ is also a map valued in $\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_{\mu}$. In other words, for any $\nu \in \mathfrak{S}_{\mu}^{\ll}$, under $\int_{\mu}^{-1}$, there is a nonempty equivalence class $[g]_{\mu}$ of Radon - Nikodym densities of ν relative to μ. We denote

$$\int_{\mu}^{-1} \text{ by the symbol } \frac{d}{d\mu}$$

and for a fixed $\nu \in \mathfrak{S}_{\mu}^{\ll}$, we set

$$\int_{\mu}^{-1}(\nu) = \frac{d\nu}{d\mu}$$

and call it the Radon - Nikodym *derivative of measure ν relative to measure μ.* We emphasize that a Radon - Nikodym derivative is not the same as a Radon - Nikodym density, but is an equivalence class of Radon - Nikodym densities. □

The following is a useful proposition and at the same time, an application of the Radon - Nikodym theorem.

8.7 Proposition. *Let μ be a σ-finite positive measure and ν a signed measure on a measurable space (Ω, Σ) such that $\nu \ll \mu$. Then there are* Radon - Nikodym *derivatives $\frac{d\nu}{d\mu}, \frac{d\nu^+}{d\mu}, \frac{d\nu^-}{d\mu}$ such that*

$$\left[\frac{d\nu}{d\mu}\right]^+ = \frac{d\nu^+}{d\mu}, \left[\frac{d\nu}{d\mu}\right]^- = \frac{d\nu^-}{d\mu},$$

In other words, given an $f \in \frac{d\nu}{d\mu}$ and its decomposition $f = f^+ - f^-$ it holds that $f^+ \in \frac{d\nu^+}{d\mu}$ and $f^- \in \frac{d\nu^-}{d\mu}$, where ν^+ and ν^- are the Jordan variations of ν.

Proof. Because $\nu \ll \mu$, by the Radon - Nikodym theorem, there is a Radon - Nikodym derivative $\frac{d\nu}{d\mu}$, which is an equivalence class $[f]_{\mu}$ of all Radon - Nikodym densities modulo μ. If $f \in \frac{d\nu}{d\mu}$, then, by Problem 6.18, the Jordan variations of ν are

$$\nu^+ = \int f^+ \, d\mu \text{ and } \nu^- = \int f^- \, d\mu \tag{8.7}$$

Equations (8.7) hold for any Radon - Nikodym densities f^+ and f^-. Without causing any confusion we write them for the entire Radon - Nikodym classes (i.e., derivatives) and we do it on occasion. So, we have (8.7) in the form

$$\nu^+ = \int \left[\frac{d\nu}{d\mu}\right]^+ d\mu \text{ and } \nu^- = \int \left[\frac{d\nu}{d\mu}\right]^- d\mu. \tag{8.7a}$$

On the other hand, by Proposition 8.2, $\nu^+ \ll \mu$ and $\nu^- \ll \mu$. Thus, by the Radon - Nikodym theorem, there are Radon - Nikodym derivatives $\frac{d\nu^+}{d\mu}$ and $\frac{d\nu^-}{d\mu}$ of ν^+ and ν^- with respect to μ. Therefore, for each $g \in \frac{d\nu^+}{d\mu}$ and $h \in \frac{d\nu^-}{d\mu}$ we have

$$\nu^+ = \int g\, d\mu \text{ and } \nu^- = \int h\, d\mu. \tag{8.7b}$$

Comparing (8.7) with (8.7b) we have

$$\int g d\mu = \int f^+ d\mu \text{ and } \int h d\mu = \int f^- d\mu,$$

where the two equations obviously hold on Σ, i.e. for each $A \in \Sigma$. By Corollary 1.21 (since μ is σ-finite),

$$g = f^+ \text{ and } h = f^- \ \mu\text{-a.e.}$$

In other words, g and f^+ and h and f^- belong to the same equivalence class, respectively, of Radon - Nikodym densities, since the latter is unique. Consequently,

$$\left[\frac{d\nu}{d\mu}\right]^+ = \frac{d\nu^+}{d\mu} \text{ and } \left[\frac{d\nu}{d\mu}\right]^- = \frac{d\nu^-}{d\mu}. \qquad \square$$

8.8 Remarks.

(i) Notice that for each $f \in \frac{d\nu}{d\mu}$, we have $f^+ \in \frac{d\nu^+}{d\mu}$ and $f^- \in \frac{d\nu^-}{d\mu}$. However, although $f = f^+ - f^-$, it is incorrect to write

$$\frac{d\nu}{d\mu} = \frac{d\nu^+}{d\mu} - \frac{d\nu^-}{d\mu}$$

or in the form

$$\frac{d\nu}{d\mu} = \left[\frac{d\nu}{d\mu}\right]^{+} - \left[\frac{d\nu}{d\mu}\right]^{-},$$

because it may cause different elements (such as f^{+} and g^{-}) of the respective classes appear in one combination. This is why most often we use two different notations for derivatives and densities.

(ii) Since ν is a signed measure, one of its Jordan variations, say ν^{+}, is finite. Consequently, $\frac{d\nu}{d\mu} \subseteq \mathbb{L}$. Thus, $\frac{d\nu^{+}}{d\mu}$ is μ-a.e. finite on Ω, whereas $\frac{d\nu^{-}}{d\mu}$ need not to. If ν is σ-finite, then by case 5b) of the Radon - Nikodym theorem, each Radon - Nikodym density of the class $\frac{d\nu}{d\mu}$ is μ-a.e. finite and so are $\frac{d\nu^{+}}{d\mu}$ and $\frac{d\nu^{-}}{d\mu}$. □

After a short break filled with various intermediate results and notions we return to a version of Radon - Nikodym's theorem for complex measures. This is readily done as follows. Firstly, given a σ-finite positive measure $\mu \in \mathfrak{M}(\Omega, \Sigma)$, we denote by

$$\mathcal{C}_{\mu}^{\ll} = \{\xi \in \mathcal{C}(\Omega, \Sigma) : \xi \ll \mu\},$$

where $\mathcal{C}(\Omega, \Sigma)$ denoted the set of all complex measures on Σ. Let $\xi \in \mathcal{C}_{\mu}^{\ll}$ and let $\xi = \nu_1 + i\nu_2$. By Proposition 8.3, $\nu_1 \ll \mu$ and $\nu_2 \ll \mu$ and ν_1 and ν_2 are finite signed measures, thus according to Case 5a of the Radon - Nikodym Theorem, there are two equivalent classes $[g_1]_{\mu}$ and $[g_2]_{\mu}$ of Radon - Nikodym densities from the factor space $L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})\big|_{\mu}$ so that, for any two elements g_1 and g_2 of their respective classes $\frac{d\nu_1}{d\mu}$ and $\frac{d\nu_2}{d\mu}$,

$$\nu_1(A) = \int_A g_1 d\mu \quad \text{and} \quad \nu_2(A) = \int_A g_2 d\mu, \quad \text{for each } A \in \Sigma,$$

thereby making $[g]_{\mu} := \frac{d\nu_1}{d\mu} \times \frac{d\nu_2}{d\mu}$ the targeted Radon - Nikodym derivative. The uniqueness of $[g]_{\mu}$ is based on that in Theorem 8.5. Summarizing the above arguments we have the following.

8.9 Theorem (Radon - Nikodym for complex measures). *Let* $\mu \in \mathfrak{M}(\Omega, \Sigma)$ *be a* σ*-finite measure. Then* $\left[L^1(\Omega, \Sigma, \mu; \mathbb{C})|_\mu, \mathcal{C}_\mu^{\ll}, \int_\mu\right]$ *is a bijective map.* □

8.10 Theorem. *Let* μ *be a* σ*-finite positive measure and* ν *be a signed measure, both on the same measurable space* (Ω, Σ) *such that* $\nu \ll \mu$. *Let* g *be any function from* $L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$. *Then, there is a unique* Radon - Nikodym *derivative class* $\frac{d\nu}{d\mu}$ *such that for any* $f \in \frac{d\nu}{d\mu}$,

$$\int g\, d\nu = \int g f d\mu.$$

Proof.

(i) Let ν be a positive measure. According to Case 4 of the Radon - Nikodym theorem, there is a nonnegative Radon - Nikodym derivative class $\frac{d\nu}{d\mu}$ such that for each $f \in \frac{d\nu}{d\mu}$, $\nu = \int f d\mu$. We can also write it in the form $\nu = \int \frac{d\nu}{d\mu} d\mu$ considering that the equation holds for each element of class $\frac{d\nu}{d\mu}$.

In other words, $\nu \ll \mu$ implies that for each measurable set A,

$$\nu(A) = \int_A f d\mu \ \ (\text{or } \nu(A) = \int_A \tfrac{d\nu}{d\mu} d\mu) \tag{8.10}$$

holding for each $f \in \frac{d\nu}{d\mu}$. We consider four cases of function g beginning with an indicator function.

(a) Let $g = \mathbf{1}_A$, for some $A \in \Sigma$. Then, by (8.10),

$$\nu(A) = \int \mathbf{1}_A d\nu = \int g d\nu = \int \mathbf{1}_A f d\mu = \int g f d\mu.$$

(b) Let g be a nonnegative simple function, $g = \sum_{i=1}^m a_i \mathbf{1}_{A_i}$. Then, by (a),

$$\begin{aligned} \int g d\nu &= \textstyle\sum_{i=1}^m a_i \int \mathbf{1}_{A_i} d\nu = \sum_{i=1}^m a_i \int \mathbf{1}_{A_i} f d\mu \\ &= \int g f d\mu. \end{aligned}$$

(c) Let $g \in \mathcal{C}_+^{-1}$. Then, there is a nondecreasing sequence $\{g_n\}$ of nonnegative simple functions convergent to g pointwise on Ω. Thus

$$\int g d\nu = \sup \int g_n d\nu = \sup \int g_n f d\mu$$

then use the monotone convergence theorem for positive measures

$$= \int g f d\mu.$$

(d) Let $g \in \mathcal{C}^{-1}(\overline{\mathbb{R}})$. Then, $g = g^+ - g^-$ and due to (c),

$$\int g d\nu = \int g^+ d\nu - \int g^- d\nu = \int g^+ f d\mu - \int g^- f d\mu$$

$$= \int (g^+ - g^-) f d\mu = \int g f d\mu.$$

The equation holds true for each $f \in \frac{d\nu}{d\mu}$. Thus we can also rewrite this equation in the form

$$\int g d\nu = \int g \tfrac{d\nu}{d\mu} d\mu.$$

(ii) Let ν be a signed measure, with the Jordan decomposition $\nu = \nu^+ - \nu^-$. Then, by Proposition 8.7, each $f \in \frac{d\nu}{d\mu}$ has the positive and negative components $f^+ \in \left[\frac{d\nu}{d\mu}\right]^+ = \frac{d\nu^+}{d\mu}$ and $f^- \in \left[\frac{d\nu}{d\mu}\right]^- = \frac{d\nu^-}{d\mu}$. Thus, for a $g \in \mathbb{L}(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$, utilizing case (i) in the integral decomposition

$$\int g d\nu = \int g d\nu^+ - \int g d\nu^-$$

separately for $\int g d\nu^+$ and $\int g d\nu^-$, we collect two densities f^+ and f^- of ν^+ and ν^- with respect to μ as two portions of f. The latter, itself, is a Radon - Nikodym density of ν with respect to μ. Thus,

$$\int g d\nu = \int g f^+ d\mu - \int g f^- d\mu = \int g (f^+ - f^-) d\mu$$

$$= \int g f d\mu.$$

The last equation holds for each density $f \in \frac{d\nu}{d\mu}$, therefore, we can rewrite it in the form

$$\int g d\nu = \int g \frac{d\nu}{d\mu} d\mu. \qquad \square$$

8.11 Theorem. *Let μ be a σ-finite positive measure and ξ be a complex measure, both on the same measurable space (Ω, Σ) such that $\xi \ll \mu$. Let g be any function from $L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}})$ or $L^1(\Omega, \Sigma; \xi, \mathbb{C})$. Then, there is a unique* Radon - Nikodym *derivative class $\frac{d\xi}{d\mu}$ such that for any $f \in \frac{d\xi}{d\mu}$,*

$$\int g \, d\xi = \int g f d\mu.$$

Proof.

(i) Now, let $\xi = \nu_1 + i\nu_2$ be a complex measure, absolutely continuous with respect to μ. Then, for a $g \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}})$,

$$\int g d\xi = \int g d\nu_1 + i \int g d\nu_2.$$

Furthermore, by the Radon - Nikodym theorem for complex measures (Theorem 8.9) we have a Radon - Nikodym density $f = u + iv \in \frac{d\xi}{d\mu}$ such that $u \in \frac{d\nu_1}{d\mu}$ and $v \in \frac{d\nu_2}{d\mu}$. Applying case (ii), we arrive at

$$\begin{aligned} \int g d\xi &= \int g u d\mu + i \int g v d\mu = \int g(u + iv) d\mu \\ &= \int g f d\mu \end{aligned}$$

or in the form

$$\int g d\xi = \int g \frac{d\xi}{d\mu} d\mu.$$

(ii) Let ξ be a complex measure such that $\xi \ll \mu$ and let $g = g_1 + ig_2 \in \mathcal{C}^{-1}(\mathbb{C})$. According to part (i) we have a Radon - Nikodym density $f \in \frac{d\nu_1}{d\mu} \times \frac{d\nu_2}{d\mu}$ such that regardless of the choice of any real-valued measurable function $h \in L^1(\Omega, \Sigma; \nu, \mathbb{R})$,

$$\int h d\xi = \int h f d\mu.$$

Thus, replacing h with g_1 or g_2 we have

$$\int (g_1 + ig_2)d\xi = \int g_1 f d\mu + i\int g_2 f d\mu = \int g f d\mu$$

or in the form

$$\int g d\xi = \int g \tfrac{d\xi}{d\mu} d\mu. \qquad \square$$

8.12 Theorem (The Chain Rule). *Let μ and ρ be two σ-finite positive measures on a measurable space (Ω, Σ) and let ξ be a signed or complex measure on (Ω, Σ). Suppose $\xi \ll \mu \ll \rho$. Then the following equation holds for* Radon - Nikodym *derivative classes*:

$$\tfrac{d\xi}{d\rho} = \tfrac{d\xi}{d\mu}\,\tfrac{d\mu}{d\rho}.$$

Proof. Suppose that $\mu \ll \rho$ such that μ and ρ are σ-finite positive measures and h is a complex function from $L^1(\Omega, \Sigma; \mu, \mathbb{C})$. Then, by Theorem 8.10, there is a density function $f \in \frac{d\mu}{d\rho}$ such that

$$\int h d\mu = \int h f d\rho. \tag{8.12}$$

or in the form of

$$\int h d\mu = \int h \tfrac{d\mu}{d\rho} d\rho. \tag{8.12a}$$

Now, if ξ is a complex measure such that $\xi \ll \mu$, then for each $h \in \frac{d\xi}{d\mu}$, (8.12-8.12a) hold, and because now $\int h d\mu = \xi$, we have

$$\xi = \int h d\mu = \int h f d\rho \tag{8.12b}$$

or in the form of

$$\xi = \int \tfrac{d\xi}{d\mu} d\mu = \int \tfrac{d\xi}{d\mu}\,\tfrac{d\mu}{d\rho} d\rho. \tag{8.12c}$$

On the other hand, by transitivity $\xi \ll \rho$, there is a $\varphi \in \frac{d\xi}{d\rho}$ such that

$$\xi = \int \varphi d\rho \tag{8.12d}$$

or in the form of

$$\xi = \int \frac{d\mu}{d\rho} d\rho. \tag{8.12e}$$

Comparing (8.12b-c) and (8.12d-e) we arrive at the equations

$$\int \varphi d\rho = \int hf d\rho \tag{8.12f}$$

or

$$\int \frac{d\mu}{d\rho} d\rho = \int \frac{d\xi}{d\mu} \frac{d\mu}{d\rho} d\rho \tag{8.12g}$$

written in terms of Radon - Nikodym derivative classes. The two integrals are integral measures that are equal on Σ and ρ is σ-finite. We remake Corollary 1.21 modifying it for complex-valued integrands φ and hf when applying it separately for the real and imaginary parts, concluding that

$$\varphi = hf \ \rho\text{-a.e. on } \Omega.$$

When equating the associated equivalence classes (i.e., Radon - Nikodym derivatives), we have the following

$$\frac{d\xi}{d\rho} = \frac{d\xi}{d\mu} \frac{d\mu}{d\rho}$$

which is reminiscent of the chain rule in calculus. □

Now, let us consider a series of applications of the Radon - Nikodym theorem in probability theory.

8.13 Example. Let X be a random variable defined on a probability space $(\Omega, \Sigma, \mathbb{P})$. Suppose $(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}), \mathbb{P}X^*)$ is the probability space induced by X, where $\mathbb{P}X^*$ is the probability distribution of X. Given a σ-finite positive Borel measure μ such that $\mathbb{P}X^* \ll \mu$, according to Case 4 of the Radon - Nikodym theorem, there is a nonempty equivalence class $\frac{d\mathbb{P}X^*}{d\mu} \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+)|_\mu$ of nonnegative Radon - Nikodym densities such that for each $g \in \frac{d\mathbb{P}X^*}{d\mu}$, it holds true that $d\mathbb{P}X^* = g d\mu$.

In particular, if $\mu = \lambda$ is the Borel - Lebesgue measure, then the probability distribution $\mathbb{P}X^*$, which is assumed to be continuous (i.e., absolutely continuous with respect to λ) can be represented by the Lebesgue integral $\int g d\lambda$, and the Radon - Nikodym density g is referred to as the *probability density function* (pdf).

Furthermore, $F(x) = \mathbb{P}X^*((-\infty, x])$ is the *probability distribution function* of X, which then can be represented by the integral $\int_{u=-\infty}^{x} g(u)\lambda(du)$. In [Dsh2] we learn that by the Lebesgue theorem, F is λ-a.e. differentiable in the usual sense and its Newton - Leibnitz derivative F' equals g λ-a.e. □

8.14 Remarks and Definition (of the Conditional Expectation).

(i) Another consequence of the Radon - Nikodym theorem, which is at the heart of stochastic analysis, is the notion of the conditional expectation. Let $X \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ be a random variable and suppose $\Sigma_0 \subseteq \Sigma$ to be a *sub-σ-algebra*. Denote $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$ and let $\nu_0 = \int X d\mathbb{P}_0$. Then, ν_0 is a signed measure on Σ_0, absolutely continuous with respect to $\mathbb{P}_0$.

Thus, according to Case 5 of the Radon - Nikodym theorem, there exists a unique Radon - Nikodym derivative (equivalence) class $\frac{d\nu_0}{d\mathbb{P}_0} \in \mathbb{L}(\Omega, \Sigma_0, \mathbb{P}_0; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))\big|_{\mu}$ of random variables (Radon - Nikodym densities), referred to as the *conditional expectation of random variable X given sub-σ-algebra Σ_0*, in notation, $\mathbb{E}[X|\Sigma_0]$. Any random variable $X_0 \in \frac{d\nu_0}{d\mathbb{P}_0}$ is called a *version of the conditional expectation*.

Now, as a Radon - Nikodym density of ν_0 with respect to $\mathbb{P}_0$, X_0 is a part of the equation $\nu_0 = \int X_0 d\mathbb{P}_0$ holding true on Σ_0. But so is X through $\nu_0 = \int X d\mathbb{P}_0$ as well, originally defined on Σ_0. Thus, we have

$$\int_A X_0 d\mathbb{P}_0 = \int_A X d\mathbb{P}_0, \text{ for all } A \in \Sigma_0. \tag{8.14}$$

In particular, with $A = \Omega$ and $\mathbb{P}_0$, lawfully replaced with $\mathbb{P}$ on the right of (8.14), we have

$$\mathbb{E}X_0 = \mathbb{E}X. \tag{8.14a}$$

(ii) In the literature on probability, $\mathbb{E}[X|\Sigma_0]$ is often confused with the above Radon - Nikodym density X_0. For notational convenience, sometimes we comply with the commonly used notation, however bearing in mind that $\mathbb{E}[X|\Sigma_0]$ is a class, not an individual Radon - Nikodym density. For example, if we replace X_0 with $\mathbb{E}[X|\Sigma_0]$ in equation (8.14a), the latter reads

$$\mathbb{E}[\mathbb{E}[X|\Sigma_0]] = \mathbb{E}X \tag{8.14b}$$

which is the well-known *law of double expectation*.

(iii) For an event (measurable set) A, let $X = \mathbf{1}_A$. Then, the conditional expectation $\mathbb{E}[\mathbf{1}_A|\Sigma_0]$ is called the *conditional probability of event A given sub-σ-algebra Σ_0,* in notation $\mathbb{P}(A|\Sigma_0)$.

(iv) If Y is a random variable on the same probability space $(\Omega, \Sigma, \mathbb{P})$, then Y is measurable with respect to the σ-algebra $\Sigma_0 := Y^{**}(\mathcal{B}(\overline{\mathbb{R}}))$ it generates in Ω, and in order to be Σ-measurable, Σ_0 must be a subset of Σ. Now, the conditional expectation of X given this sub-σ-algebra in probability theory is often denoted by $\mathbb{E}[X|Y]$. □

The following construction explains why $\mathbb{E}[X|\Sigma_0]$ is called the "conditional expectation."

8.15 Examples.

(i) Let X be a random variable on a probability space $(\Omega, \Sigma, \mathbb{P})$ and let $\mathbf{H} = \{H_1, H_2, \dots\}$ be a measurable partition of Ω such that $\mathbb{P}(H_n) > 0$ for each $n \in \mathbb{N}$. Then,

$$\Sigma_0 = \left\{\Omega, \emptyset, A = \sum_{i\in J} H_i : J \subseteq \mathbb{N}\right\} \tag{8.15}$$

is a σ-algebra generated by $\mathbf{H}$. Clearly, $\Sigma_0 \subseteq \Sigma$. Let $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$ and define the random variable X_0 on the probability space $(\Omega, \Sigma_0, \mathbb{P}_0)$:

$$X_0 = \sum_{n=1}^{\infty} \left(\frac{1}{\mathbb{P}(H_n)} \int_{H_n} X d\mathbb{P}_0 \right) \mathbf{1}_{H_n}. \tag{8.15a}$$

For

$$X_0 \mathbf{1}_{H_k} = \left(\frac{1}{\mathbb{P}(H_k)} \int_{H_k} X d\mathbb{P}_0 \right) \mathbf{1}_{H_k},$$

we have

$$\mathbb{E}[X_0 \mathbf{1}_{H_k}] = \int_{H_k} X_0 d\mathbb{P}_0 = \int_{H_k} X d\mathbb{P}_0 = \mathbb{E}[X \mathbf{1}_{H_k}],\ k = 1, 2, \dots$$

Next, for an $A \in \Sigma_0$, we calculate for

$$X_0 \mathbf{1}_A = X_0 \mathbf{1}_{\sum_{i \in J} H_i} = \sum_{n \in J} \left(\frac{1}{\mathbb{P}(H_n)} \int_{H_n} X d\mathbb{P}_0 \right) \mathbf{1}_{H_n}$$

the expected value

$$\mathbb{E}[X_0 \mathbf{1}_A] = \sum_{n \in J} \int_{H_n} X d\mathbb{P}_0 = \int_A X d\mathbb{P}_0 = \mathbb{E}[X \mathbf{1}_A].$$

Hence,

$$\nu_0 = \int X d\mathbb{P}_0 = \int X_0 d\mathbb{P}_0 \ \text{ on } \Sigma_0$$

and therefore, by the Radon - Nikodym theorem, X_0 is a version of the conditional expectation $\mathbb{E}[X|\Sigma_0]$. In fact, in this particular, case, this is the only version of $\mathbb{E}[X|\Sigma_0]$. This is because the σ-algebra Σ_0 is too crude and it does not contain any null sets other than the empty set. Suppose N is a $\mathbb{P}$-null set, not listed in Σ_0, and let

$$\widetilde{X} = X_0 + \mathbf{1}_N.$$

Then, obviously, $\widetilde{X} = X_0$ $\mathbb{P}$-a.s. However, $\widetilde{X}$ is not a version of $\mathbb{E}[X|\Sigma_0]$, because $\widetilde{X}$ is not Σ_0-measurable to begin with.

(ii) Now, let us include N in Σ_0. To simplify the construction, assume that $\mathrm{H} = \{H_0, H_1, \dots\}$ is a measurable partition of Ω such

that $H_0 = N$ and thus $\mathbb{P}(H_0) = 0$, whereas $\mathbb{P}(H_n) > 0, n \in \mathbb{N}$. We modify Σ_0 in (8.15) as

$$\Sigma_0 = \left\{ \Omega, \emptyset, A = \sum_{i \in J} H_i : J \subseteq \mathbb{N}_0 \right\} \tag{8.15b}$$

and define X_0 exactly as in (8.15a). Now, obviously,

$$\mathbb{E}[X_0 \mathbf{1}_N] = \mathbb{E}[X \mathbf{1}_N] = 0$$

and therefore, again we can show that

$$\nu_0 = \int X d\mathbb{P}_0 = \int X_0 d\mathbb{P}_0 \text{ on } \Sigma_0.$$

However, in this case $\widetilde{X} = X_0 + \mathbf{1}_N$ is Σ_0-measurable and $\widetilde{X}$ turns out to be yet another version of the conditional expectation $\mathbb{E}[X|\Sigma_0]$.

(iii) In the context of the above constructions, let B and H be two events such that $\mathbb{P}(H) > 0$. We define the conditional expectation

$$\mathbb{E}[X|H] = \tfrac{1}{\mathbb{P}(H)} \int\limits_H X d\mathbb{P}. \tag{8.15c}$$

In particular, for the random variable $\mathbf{1}_B$ we have

$$\mathbb{E}[\mathbf{1}_B|H] = \tfrac{\mathbb{P}(B \cap H)}{\mathbb{P}(H)}. \tag{8.15d}$$

The latter agrees with the conditional probability $\mathbb{P}(B|H)$. Now, if in (8.15a) we replace $\frac{1}{\mathbb{P}(H_n)} \int\limits_{H_n} X d\mathbb{P}_0$ with $\mathbb{E}[X|H_n]$ as per (8.15c), we can rewrite (8.15a) as

$$X_0 = \textstyle\sum_{n=1}^{\infty} \mathbb{E}[X|H_n] \mathbf{1}_{H_n}. \tag{8.15e}$$

X_0 is the only version of the conditional expectation, thus we can also write

$$\mathbb{E}[X|\Sigma_0] = \textstyle\sum_{n=1}^{\infty} \mathbb{E}[X|H_n] \mathbf{1}_{H_n}. \tag{8.15f}$$

□

In closing on applications of the Radon - Nikodym theorem to probability we offer several useful propositions that the reader can prove with relatively little effort.

8.16 Proposition. *Let* $X \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ *be a random variable*, $\Sigma_0 \subseteq \Sigma$ *a sub-σ-algebra, and* $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. *Then,* $X = \mathbb{E}[X|\Sigma_0]$ $\mathbb{P}_0$*-a.s. if* X *is* Σ_0*-measurable*.

8.17 Proposition (linearity of the conditional expectation). *Let* $X, Y \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-σ-algebra, and* $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. *Then, for any two real numbers* a *and* b,

$$\mathbb{E}[aX + bY|\Sigma_0] = a\mathbb{E}[X|\Sigma_0] + b\mathbb{E}[Y|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

8.18 Proposition (monotonicity of the conditional expectation). *Let* $X, Y \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-σ-algebra, and* $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. *Then, if* $X \leq Y$ $\mathbb{P}$*-a.s.,*

$$\mathbb{E}[X|\Sigma_0] \leq \mathbb{E}[Y|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

8.19 Proposition (monotone convergence theorem for the conditional expectation). *Let* $X_1, X_2, \ldots$ *be a monotone nondecreasing sequence of nonnegative r.v.'s from* $\mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-σ-algebra, and* $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. *Then*

$$\lim_{n\to\infty}\mathbb{E}[X_n|\Sigma_0] = \mathbb{E}[\lim_{n\to\infty}X_n|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

8.20 Proposition. *Let* $X, Y,$ *and* $XY \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-σ-algebra, and* $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. *Then, if* Y *is* Σ_0*-measurable*,

$$\mathbb{E}[XY|\Sigma_0] = Y\mathbb{E}[X|\Sigma_0] \ \mathbb{P}_0\text{-a.s.} \qquad \square$$

PROBLEMS

8.1 Prove Proposition 8.2. *Let ν be a signed measure and μ - a positive measure on a measurable space (Ω, Σ). Then the following statements hold*:

(i) $\nu \ll \mu \Rightarrow \nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow |\nu| \ll \mu$.

(ii) $\nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow \nu \ll \mu$ and $|\nu| \ll \mu$.

(iii) $|\nu| \ll \mu \Rightarrow \nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow \nu \ll \mu$.

8.2 Prove Proposition 8.3. *Let $\xi = \nu + i\rho$ be a complex measure and μ a positive measure on a measurable space (Ω, Σ). Then, the following statements are equivalent*:

(i) $\xi \ll \mu$.

(ii) $|\xi| \ll \mu$.

(iii) $\nu^+ \ll \mu$, $\nu^- \ll \mu$, $\rho^+ \ll \mu$, $\rho^- \ll \mu, \nu \ll \mu$, $\rho \ll \mu$.

8.3 Let (Ω, Σ, ξ) be a complex measure space. Show that $\xi \ll |\xi|$.

8.4 Let $\xi \in \mathcal{C}(\Omega, \Sigma)$. Show that the Radon - Nikodym derivative $\frac{d\xi}{d|\xi|}$ of ξ with respect to $|\xi|$ exists and that the modulus of any Radon - Nikodym density from $\frac{d\xi}{d|\xi|}$ equals $\mathbf{1}$ $|\xi|$-a.e., where $\mathbf{1}$ is the function defined as $\mathbf{1}(\omega) = 1, \omega \in \Omega$.

8.5 Consider in Case 1 of the Radon - Nikodym theorem, the partial order $\preceq$ on Φ by defining $f \preceq g$ if and only if $f \leq g$ μ-a.e. Show that any chain in Φ has an upper bound and thus, by Zorn's Lemma 4.13, Chapter 1, Φ has a maximal element.

8.6 Let $\xi \in \mathcal{C}(\Omega, \Sigma)$. Show that the Radon - Nikodym derivative $\frac{d\xi}{d|\xi|}$ exists and that for each $g \in L^1(\Omega, \Sigma; \xi, \mathbb{C})$,

$$\int g d\xi = \int g \tfrac{d\xi}{d|\xi|}\, d|\xi|.$$

8.7 Let μ and ρ be two σ-finite signed measures on a measurable space (Ω, Σ) such that $\mu \approx \rho$ (meaning $\mu \ll \rho$ and $\rho \ll \mu$). Show that

$$\tfrac{d\rho}{d\mu} = \left(\tfrac{d\mu}{d\rho}\right)^{-1} \mu\text{- (or } \rho\text{-) a.e.}$$

8.8 Show that $\mathcal{C}_\mu^{\ll}$ is a linear space over the field $\mathbb{C}$. Does the same apply to $\mathfrak{S}_\mu^{\ll}$?

8.9 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space under the following assumptions. Let $\Omega = [0,1]$, $\Sigma = \mathcal{B} \cap [0,1]$, and $\mathbb{P} = \mathrm{Res}_\Sigma \lambda$ (where λ denotes the Borel - Lebesgue measure). Let $\mathbf{H} = \{H_1, \ldots, H_n\}$ be a finite partition of Ω such that

$$H_k = \left[\tfrac{k-1}{n}, \tfrac{k}{n}\right), k = 1, \ldots, n,$$

and let Σ_0 be the smallest σ-algebra generated by $\mathbf{H}$. Let $X(\omega) = \omega$, for all $\omega \in \Omega$. Calculate a version X_0 of the conditional expectation $\mathbb{E}[X|\Sigma_0]$ and give all other versions of $\mathbb{E}[X|\Sigma_0]$.

8.10 Prove Proposition 8.16. *Let $X \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ be a random variable, $\Sigma_0 \subseteq \Sigma$ a sub-σ-algebra, and $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. Then, $X = \mathbb{E}[X|\Sigma_0]$ $\mathbb{P}_0$-a.s. if X is Σ_0-measurable.*

8.11 Explain exactly why $X \neq \mathbb{E}[X|\Sigma_0]$ if X is not Σ_0-measurable.

8.12 Let X be a random variable on a probability space $(\Omega, \Sigma, \mathbb{P})$ such that $X = \mathbf{1}$ $\mathbb{P}$-a.s. on Ω. Let Σ_0 be a sub-σ-algebra of Σ on Ω and $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. Show that $X = E[X|\Sigma_0]$ $\mathbb{P}_0$-a.s.

8.13 Prove Proposition 8.17. *Let* $X, Y \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-*σ*-algebra, and* $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. *Then, for any two real numbers* a *and* b,

$$\mathbb{E}[aX + bY|\Sigma_0] = a\mathbb{E}[X|\Sigma_0] + b\mathbb{E}[Y|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

8.14 Prove Proposition 8.18. *Let* $X, Y \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-*σ*-algebra, and* $\mathbb{P}_0 = Res_{\Sigma_0}\mathbb{P}$. *Then, if* $X \leq Y$ $\mathbb{P}$-a.s.,

$$\mathbb{E}[X|\Sigma_0] \leq \mathbb{E}[Y|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

8.15 Prove Proposition 8.19. *Let* $X_1, X_2, \ldots$ *be a monotone non-decreasing sequence of nonnegative r.v.'s from* $\mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-*σ*-algebra, and* $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. *Then*

$$\lim_{n\to\infty}\mathbb{E}[X_n|\Sigma_0] = \mathbb{E}[\lim_{n\to\infty}X_n|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

8.16 Prove Proposition 8.20. *Let* $X, Y,$ *and* $XY \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-*σ*-algebra, and* $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. *Then, if* Y *is* Σ_0*-measurable*,

$$\mathbb{E}[XY|\Sigma_0] = Y\mathbb{E}[X|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

NEW TERMS:

absolutely continuous signed or complex measure with respect to a positive measure, 471
absolutely continuous signed or complex measure with respect to a signed or complex measure, 471
$\mathfrak{S}_{\mu}^{\ll} = \{\nu \in \mathfrak{S}(\Omega, \Sigma) : \nu \ll \mu\}$, set of all signed measures continuous relative to μ, 472
$[\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_{\mu}, \mathfrak{S}_{\mu}^{\ll}, \int_{\mu}]$, bijective map between $\mathbb{L}$ functions and signed measures continuous relative to μ, 472
Radon - Nikodym density of a signed measure, 473
Radon - Nikodym theorem for a signed measure, 473
Radon - Nikodym theorem for two finite positive measures, 474
Radon - Nikodym theorem for μ finite and ν σ-finite positive measures, 477
Radon - Nikodym theorem for μ finite and ν arbitrary positive measures, 477
Radon - Nikodym theorem for μ σ-finite and ν arbitrary positive measures, 479
Radon - Nikodym theorem for μ σ-finite and ν arbitrary signed measures, 479
$\int_{\mu}^{-1}(\nu) = \frac{d\nu}{d\mu}$, Radon - Nikodym derivative of a signed measure, 481
$\frac{d\nu}{d\mu}, \frac{d\nu^{+}}{d\mu}, \frac{d\nu^{-}}{d\mu}$, Radon - Nikodym derivatives of signed measure ν and Jordan variations with respect to μ, 481
$\mathcal{C}_{\mu}^{\ll} = \{\xi \in \mathcal{C}(\Omega, \Sigma) : \xi \ll \mu\}$, set of all complex measures continuous relative to μ, 483
Radon - Nikodym theorem for complex measures, 484
$\left[L^{1}(\Omega, \Sigma, \mu; \mathbb{C})|_{\mu}, \mathcal{C}_{\mu}^{\ll}, \int_{\mu}\right]$, bijective map between L^{1} functions and signed measures continuous relative to μ, 484
the chain rule, 487
probability density and probability distribution functions, 489
conditional expectation given a sub-σ-algebra, 489
the law of double expectation, 490
conditional probability of an event given a sub-σ-algebra, 490
linearity and monotonicity of the conditional expectation, 493
monotone convergence theorem for the conditional expectation, 493

9. MEASURES AND INTEGRATION ON FINITE PRODUCT SPACES

The present section extends the results on integration to Cartesian products. It discusses the formation of product σ-algebras (which has some resemblance to the product topology) and product measures on them. This leads to the main result of this section, the celebrated Fubini's theorem, which allows one to iterate multiple integrals as its measure-theoretic analogue of multiple Riemann integrals.

Many textbooks in analysis and on the history of mathematics adopt "Fubini's theorem" as a generic name for a class of theorems establishing the identity of multiple integrals with iterated integrals. In the mainstream of the evolution of calculus, when integrating a function f on the rectangle $R = [a, b] \times [c, d]$, the question was raised: under what condition does the existence of the double integral $\int\int_R f d(x, y)$ guarantee the existence of either of the iterated integrals, $\int_a^b \{\int_c^d f(x, y)dy\}dx$ and $\int_c^d \{\int_a^b f(x, y)dx\}dy$, and will they all be equal? Fubini's theorem, in one of its earlier forms, was proved by Augustin-Louis Cauchy in the early nineteenth century and applied to continuous functions. In 1904, Henri Lebesgue extended this result to bounded measurable functions. In 1906 Beppo Levi conjectured that f need not be bounded, but just integrable. The Italian Guido Fubini (1879 – 1943) proved this statement in 1907. Namely, he proved that given the function f is integrable on R, the functions $x \mapsto f(x, y)$ and $y \mapsto f(x, y)$ are integrable for almost all x and y, respectively. In addition, the functions $y \mapsto \int_a^b f(x, y)dx$ and $x \mapsto \int_c^d f(x, y)dy$ are integrable and

$$\int\int_R f d(x, y) = \int_a^b dx\{\int_c^d f(x, y)dy\} = \int_c^d dy\{\int_a^b f(x, y)dx\}.$$

Fubini, however, imposed an unnecessary condition on the integrand function. This was corrected and refined independently by the Italian Leonida Tonelli (1885 – 1946) in 1909, Brit Ernest W. Hobson (1856 – 1933), and Belgian Charles J.G.N. (Baron) de la Valée-

Poussin (1866 – 1962) in 1910 who rendered their proofs entirely differently from that of Fubini.

The notion of multiple integrals goes back to as early as the middle of the eighteenth century, first in the form of an indefinite integral. Later on, by 1770, Leonard Euler formalized the double integral on a bounded domain and applied the above formula for iterated integrals by justifying it in terms of Riemann sums. Functions to be integrated were assumed to be continuous, and the area of integration was fairly basic. This approach began to encounter serious conflicts as soon as more general cases were considered. Not until Henri Lebesgue published his famous thesis in 1902, did it become possible to tackle other classes of functions, which all led to Fubini's theorem of its 1910 version. More general versions of Fubini's theorem (which we explore in this section) in abstract spaces, applied to abstract measures and integrals, became accomplishable after the Johann Radon's extension of the Lebesgue integral in 1913 (mentioned in Section 5).

We begin this section with the notion of "product σ-algebra" which has many similarities with the formation of a product topology with finitely many factor spaces. As in the event of a product topology, we define the product σ-algebra on the space of rectangles and show that it is identical to the σ-algebra relative to which all projection maps are measurable (continuous in the case of the topology).

9.1 Definition. Let $(\Omega_i, \Sigma_i),\ i = 1, \ldots, n,$ be measurable spaces. Denote

$$\overset{n}{\underset{i=1}{\times}} \Sigma_i = \left\{ \prod_{i=1}^{n} A_i : A_i \in \Sigma_i, i = 1, \ldots, n \right\} \tag{9.1}$$

and call it the set of all *measurable rectangles* in $\Omega := \prod_{i=1}^{n} \Omega_i$. The σ-algebra on Ω generated by $\overset{n}{\underset{i=1}{\times}} \Sigma_i$ is called the *product σ-algebra* and is denoted by

$$\Sigma_{\otimes} = \overset{n}{\underset{i=1}{\otimes}} \Sigma_i = \Sigma\Big(\overset{n}{\underset{i=1}{\times}} \Sigma_i \Big). \tag{9.1a}$$

□

9.2 Theorem. *$\Sigma_\otimes$ is the smallest σ-algebra on Ω relative to which all projection maps are measurable, that is, for each $i = 1, \dots, n$, π_i is $\Sigma_\otimes$-Σ_i-measurable. In other words,*

$$\Sigma_\otimes = \Sigma\left(\bigcup_{i=1}^{n} \pi_i^{**}(\Sigma_i)\right). \tag{9.2}$$

Proof. (i) Let $\mathfrak{A}$ be a σ-algebra relative to which all projection maps are measurable. In other words, $\pi_i^{**}(\Sigma_i) \subseteq \mathfrak{A}$. We show that $\overset{n}{\underset{i=1}{\times}} \Sigma_i \subseteq \mathfrak{A}$. Indeed, if R is a measurable rectangle (i.e. $R \in \overset{n}{\underset{i=1}{\times}} \Sigma_i$), $R = \bigcap_{i=1}^{n} \pi_i^*(A_i)$ such that $A_i \in \Sigma_i$ and since $\pi_i^*(A_i) \in \mathfrak{A}$, $R \in \mathfrak{A}$.

(ii) Because $\overset{n}{\underset{i=1}{\times}} \Sigma_i$ is a generator of $\Sigma_\otimes$, we have $\Sigma_\otimes \subseteq \mathfrak{A}$.

(iii) We show that π_i is $\Sigma_\otimes$-Σ-measurable. Indeed, if $A_i \in \Sigma_i$, the measurable cylinder $\pi_i^*(A_i) \in \overset{n}{\underset{i=1}{\times}} \Sigma_i$, because it also is a measurable rectangle. Thus, $\pi_i^{**}(\Sigma_i) \subseteq \overset{n}{\underset{i=1}{\times}} \Sigma_i \subseteq \Sigma_\otimes$.

(iv) In conclusion, because $\Sigma_\otimes \subseteq \mathfrak{A}$, and since $\mathfrak{A}$ is a σ-algebra relative to which all projections are measurable, $\Sigma_\otimes$ is the smallest of all such σ-algebras. On the other hand, $\Sigma\left(\bigcup_{i=1}^{n} \pi_i^{**}(\Sigma_i)\right)$ is yet another smallest σalgebra relative to which all projections are measurable. □

Consequently, we say that $\Sigma_\otimes$ is *generated by the projection maps*. We recollect that the Borel σ-algebra on $\mathbb{R}^n$ was generated half-open n-dimensional intervals (called *rectangles*) and that the collection of all such rectangles was a semi-ring. It seems plausible that $\overset{n}{\underset{i=1}{\times}} \Sigma_i$ is also a semi-ring.

9.3 Lemma. *Let $A, B \in \overset{n}{\underset{i=1}{\times}} \Sigma_i$. Then, $A \backslash B$ can be partitioned with a finite union of pairwise disjoint measurable rectangles.*

Proof. (i) For $n = 2$, we have from Problem 5.9, Chapter 1, that

$$A\backslash B = [(A_1\backslash B_1) \times A_2] + [(A_1 \cap B_1) \times (A_2\backslash B_2)].$$

According to the same problem, there is yet another, alternative, decomposition of $A\backslash B$:

$$A\backslash B = [A_1 \times (A_2\backslash B_2)] + [(A_1\backslash B_1) \times (A_2 \cap B_2)].$$

(ii) For an arbitrary n, we proceed by induction. Thus, we assume that if $A, B \in \underset{i=1}{\overset{n}{\times}} \Sigma_i$, $A\backslash B = \sum_{s=1}^{n} R_s^{(n)}$, where $R_s^{(n)} \in \underset{i=1}{\overset{n}{\times}} \Sigma_i$. Now consider

$$\underset{i=1}{\overset{n+1}{\times}} \Sigma_i = \underset{i=1}{\overset{n}{\times}} \Sigma_i \times \Sigma_{n+1} = : \Sigma^{(1)} \times \Sigma^{(2)}.$$

Let $A, B \in \Sigma^{(1)} \times \Sigma^{(2)}$. Thus, $A = A^{(1)} \times A^{(2)}$ and $B = B^{(1)} \times B^{(2)}$. By Case (i),

$$A\backslash B = \left[\left(A^{(1)}\backslash B^{(1)}\right) \times A^{(2)}\right] + \left[\left(A^{(1)} \cap B^{(1)}\right) \times \left(A^{(2)}\backslash B^{(2)}\right)\right]$$

$$= \left[\left(\sum_{s=1}^{n} R_s^{(n)}\right) \times A^{(2)}\right] + \left[R_{s+1}^{(n)} \times \left(A^{(2)}\backslash B^{(2)}\right)\right]$$

(by Problem 5.6(i), Chapter 1)

$$= \sum_{s=1}^{n} \left(R_s^{(n)} \times A^{(2)}\right) + R_{s+1}^{(n+1)} = \sum_{s=1}^{n+1} R_s^{(n+1)}.$$

Now, because $\underset{i=1}{\overset{n}{\times}} \Sigma_i$ is $\cap$-stable, it is a semi-ring. □

9.4 Remark. It seems plausible to define an elementary content μ^0 on $\underset{i=1}{\overset{n}{\times}} \Sigma_i$ by setting $\mu^0(A) = \prod_{i=1}^{n} \mu_i(A_i)$ analogous to our construction of the Lebesgue elementary content on $\mathbb{R}^n$. Using Lemma 2.5 and Theorem 2.6 of Chapter 4, we can extend the semi-ring $\underset{i=1}{\overset{n}{\times}} \Sigma_i$ to the smallest ring (in our case it is an algebra),

$\mathcal{A}\Big(\overset{n}{\underset{i=1}{\times}}\Sigma_i\Big)$ of measurable elements (called *figures*), each one of which is either a rectangle or it can be partitioned in finitely many disjoint rectangles. Consequently, μ^0 can be extended to a content $\widetilde{\mu}$ to measure any figure $F \in \mathcal{A}\Big(\overset{n}{\underset{i=1}{\times}}\Sigma_i\Big)$ as, e.g. $\widetilde{\mu}(F) = \sum_{s=1}^{m}\mu^0(R_s)$. We can go on in the direction of Carathéodory extension, making some additional assumptions. However, we will render the extension another way by specifying an integral measure (fully explicitly), which operates on $\Sigma_\otimes$ so that it equals μ^0 on $\overset{n}{\underset{i=1}{\times}}\Sigma_i$. □

Until we get to that point, for the sequel, we establish some auxiliary results. The lemma below enables us to utilize a more economical generator of $\Sigma_\otimes$, again with some resemblance to the formation of the product topology. Consider one more notation. Let $\mathcal{G}_i$ be a subfamily of measurable sets from Σ_i. Let us denote by

$$\mathcal{G} = \overset{n}{\underset{i\,=\,1}{\times}}\mathcal{G}_i = \mathcal{G}_1 \times \cdots \times \mathcal{G}_n$$

the set of all measurable rectangles $G_1 \times \cdots \times G_n$ where G_i's are picked from $\mathcal{G}_i$.

9.5 Lemma. *Let $\mathcal{G}_i$ be a generator of Σ_i containing a sequence $\{G_{ik}\colon k = 1, 2, \dots\}$ of sets, $i = 1, \dots, n$, monotonically increasing to Ω_i. Then the product σ-algebra $\Sigma_\otimes$ coincides with the σ-algebra*

$$\Sigma = \Sigma\Big(\overset{n}{\underset{i\,=\,1}{\times}}\mathcal{G}_i\Big) \text{ (generated by } \mathcal{G} = \overset{n}{\underset{i\,=\,1}{\times}}\mathcal{G}_i).$$

Proof.

(i) Because $\mathcal{G} \subseteq \Sigma_\otimes$, it follows that $\Sigma \subseteq \Sigma_\otimes$. Indeed, every $A_i \in \mathcal{G}_i$ is also an element of Σ_i and therefore $\pi_i^*(A_i) \in \pi_i^*(\Sigma_i) \subseteq \Sigma$ which implies that

$$\prod_{i=1}^{n} A_i = \bigcap_{i=1}^{n} \pi_i^*(A_i) \in \Sigma_\otimes.$$

(ii) Now we show the inverse inclusion $\Sigma_\otimes \subseteq \Sigma$. We prove that each π_i is Σ-Σ_i-measurable. By our assumption, each generator of Σ_i contains a sequence $\{G_{in}\} \uparrow \Omega_i$. Consider

$$G_{1k} \times G_{2k} \times \cdots \times G_{i-1,k} \times A_i \times G_{i+1,k} \times \cdots \times G_{n,k}$$

$$= \widetilde{G}_k \in \mathcal{G},$$

where $A_i \in \mathcal{G}_i$. Observe that

$$\widetilde{G}_k = \pi_1^*(G_{1k}) \cap \cdots \cap \pi_i^*(A_i) \cap \cdots \cap \pi_n^*(G_{nk}).$$

Therefore,

$$\sup_k\{\widetilde{G}_k\} = \bigcup_{k=1}^{\infty} \widetilde{G}_k = \pi_i^*(A_i) \in \Sigma$$

(because we took the union of elements of Σ). Hence, we proved that the inverse image of an arbitrary element of $\mathcal{G}_i$ (which is a generator of Σ_i) under π_i belongs to Σ. According to Proposition 3.4, Chapter 4, we claim that the same inclusion holds for an arbitrary element of Σ_i or that π_i is Σ-Σ_i-measurable. In as much as π_i is $\Sigma_\otimes$-Σ_i-measurable for all $i = 1, \ldots, n$, it follows that Σ contains all measurable cylinders, that is, a generator of $\Sigma_\otimes$. Thus Σ contains $\Sigma_\otimes$. □

9.6 Remark. Now we see that in light of Lemma 9.5, $\Sigma_\otimes = \bigotimes_{i=1}^{n} \Sigma_i$ is induced by a more "economical" generator than the one given in Definition 9.1, i.e. by all rectangles from Σ_i's. In some cases, when we fail to indicate a generator, we do consider $\Sigma_\otimes$ as the σ-algebra generated by all rectangles as per Lemma 9.5. □

9.7 Examples. To tell sets and set collections in $\mathbb{R}$ from those in $\mathbb{R}^n$ we will attach to the latter the superscript ^{n}.

(i) Let $\Omega_i = \mathbb{R}$, $\Sigma_i = \mathcal{B}_i = \mathcal{B}$ (Borel σ-algebra), $i = 1, \ldots, n$. Then $\Omega = \mathbb{R}^n$ and $\mathcal{B}_\otimes = \bigotimes_{i=1}^{n} \mathcal{B}_i$. We also know that there is another σ-algebra in $\mathbb{R}^n$, the Borel σ-algebra $\mathcal{B}^n = \mathcal{B}(\mathbb{R}^n)$. What is the relationship between $\mathcal{B}^n$ and $\mathcal{B}_\otimes$? Recall that $\mathcal{B}^n$ was generated by the semi-ring of n-dimensional half-open intervals, which, for convenience, we denote by $\mathcal{S}^n$. Observe that

$$\mathcal{S}^n = \underbrace{\mathcal{S} \times \cdots \times \mathcal{S}}_{n}$$

and that each $\mathcal{S}$ contains a sequence, monotonically increasing to $\mathbb{R}$. Thus, by Lemma 9.5, $\mathcal{B}^n = \Sigma(\mathcal{S}_n)$ and $\mathcal{B}_\otimes$ must coincide.

(ii) Recall that the Borel - Lebesgue measure λ^n on $\mathcal{B}^n$ was emerged from the Lebesgue elementary content λ^{0n}, defined on $\mathcal{S}^n$ as

$$\lambda^{0n}((a^n, b^n]) = \prod_{i=1}^{n}(b_i - a_i) = \prod_{i=1}^{n}\lambda_0((a_i, b_i]).$$

We know that λ^n is a unique extension of that content on $\mathcal{B}_\otimes$. In light of Remark 9.4, we can look at a more general problem. Let us now consider an n-tuple of measure spaces $(\Omega_i, \Sigma_i, \mu_i)$, $i = 1, \ldots, n$. We wonder if there exists a unique measure μ on the measurable space $\left(\prod_{i=1}^{n}\Omega_i\,,\ \bigotimes_{i=1}^{n}\Sigma_i\right)$ such that for each measurable rectangle A,

$$\mu(A) = \mu(A_1 \times \cdots \times A_n) = \mu_1(A_1)\cdots\mu_n(A_n). \qquad \square$$

We will begin with a construction of such a measure to be referred as a *product measure*.

9.8 Definition. Let B be an arbitrary subset of $\Omega = \prod_{i=1}^{n}\Omega_i$. We define for a point $a_i \in \Omega_i$ the a_i*-section* of B as

$$\begin{aligned} B_{a_i} = \{&(\omega_1, \ldots, \omega_{i-1}, \omega_{i+1}, \ldots, \omega_n) \\ &\in \Omega_1 \times \cdots \times \Omega_{i-1} \times \Omega_{i+1} \times \cdots \times \Omega_n: \\ &(\omega_1, \ldots, \omega_{i-1}, a_i, \omega_{i+1}, \ldots, \omega_n) \in B\}. \end{aligned} \tag{9.8}$$

If a_i is such that $a_i \notin \pi_{i*}(B)$, then $(\omega_1, \ldots, \omega_{i-1}, a_i, \omega_{i+1}, \ldots, \omega_n) \notin B$ and $B_{a_i} = \emptyset$. (See Figure 9.1 below.)

Here π_{ij} was defined as $\pi_{ij} : \Omega \to \Omega_i \times \Omega_j$.

For the case of two factor spaces, for each $a_i \in \Omega_i$ we can formalize a_i*-sections* as maps $[\Sigma_\otimes, \mathcal{P}(\Omega_2), \sigma_{a_1}]$ and $[\Sigma_\otimes, \mathcal{P}(\Omega_1), \sigma_{a_2}]$ such that for any $B \in \Sigma_\otimes$,

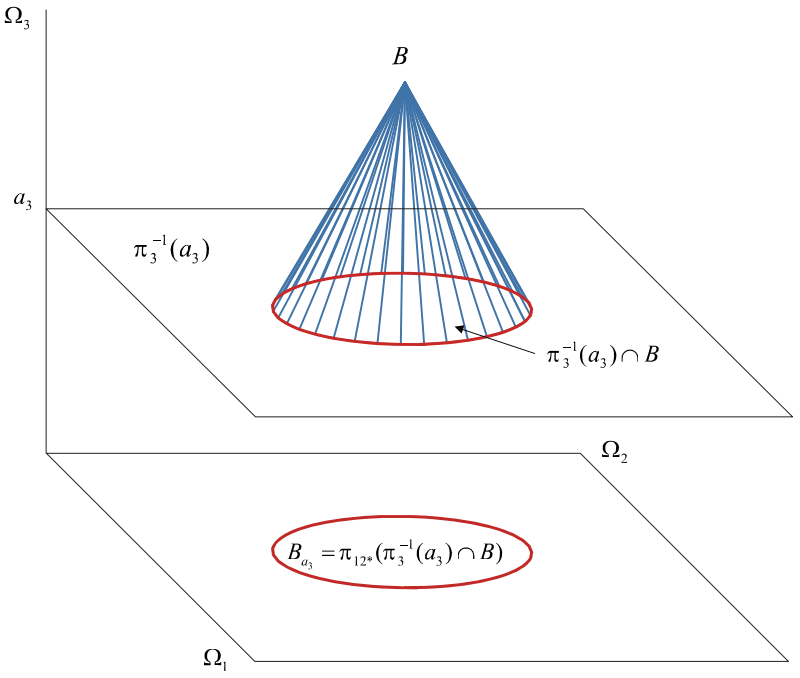


Figure 9.1

$$B_{a_1} = \sigma_{a_1}(B)$$

$$= \begin{cases} \{y \in \Omega_2 : (a_1, y) \in B\}, & a_1 \in \pi_{1*}(B) \\ \text{Ø}, & a_1 \notin \pi_{1*}(B) \end{cases} \tag{9.8a}$$

and

$$B_{a_2} = \sigma_{a_2}(B)$$

$$= \begin{cases} \{x \in \Omega_1 : (x, a_2) \in B\}, & a_2 \in \pi_{2*}(B) \\ \text{Ø}, & a_2 \notin \pi_{2*}(B). \end{cases} \tag{9.8b}$$

□

Without loss of generality we continue to work with two factor spaces. The lemma below claims that the codomains of the above maps in (9.8a - b) can be narrowed down to Σ_2 and Σ_1, respectively.

9.9 Lemma. *Let σ_{a_i} be a_i-sections for some $a_i \in \Omega_i$, $i = 1, 2$. Then they are the maps $[\Sigma_\otimes, \Sigma_2, \sigma_{a_1}]$ and $[\Sigma_\otimes, \Sigma_1, \sigma_{a_2}]$. In other words, for each $A \in \Sigma_\otimes$, $\sigma_{a_1}(A)$ is Σ_2-measurable and $\sigma_{a_2}(A)$ is Σ_1-measurable.*

Proof. Denote $\Sigma' = \{A \in \Sigma_\otimes\colon\ \sigma_{a_1}(A) = A_{a_1} \in \Sigma_2\}$. We need to prove statements (a) and (b).

(a) As per Problem 9.4, Σ' is a σ-algebra in $\Omega_1 \otimes \Omega_2$. The earlier Problem 9.3 claims that the section is commutative with all set operations which helps in proving assertion (a).

(b) Any measurable rectangle is an element of Σ'. This assertion is based on the property of the section

$$\sigma_{a_1}(A_1 \times A_2) = (A_1 \times A_2)_{a_1} = \begin{cases} A_2, & a_1 \in A_1 \\ \varnothing, & a_1 \notin A_1 \end{cases} \tag{9.9}$$

(which is easy to prove using the definition of σ_{a_1}), thereby making $\Sigma_1 \times \Sigma_2$ a subset of Σ'. By (a), Σ' is a σ-algebra, and the set of all measurable rectangles $\Sigma_1 \times \Sigma_2$ is a generator of $\Sigma_\otimes$, therefore it follows that the latter is contained in Σ'. On the other hand, due to the definition of Σ', $\Sigma' \subseteq \Sigma_\otimes$. Therefore, $\Sigma' = \Sigma_\otimes$ and concludes the claim that $\sigma_{a_1 *}(\Sigma_\otimes) \subseteq \Sigma_2$. Of course, the proof that $\sigma_{a_2 *}(\Sigma_\otimes) \subseteq \Sigma_1$ is identical. □

We return to the problem of formalizing a product measure on $\Sigma_\otimes$ posed in Remark 9.4.

If μ_i is a measure on Σ_i, $i = 1, 2$, then according to Lemma 9.9, for an arbitrary set $A \in \Sigma_1 \otimes \Sigma_2$, $A_{a_1} \in \Sigma_2$ and $A_{a_2} \in \Sigma_1$, and thus $\mu_1(A_{a_2})$ and $\mu_2(A_{a_1})$ are defined terms. For a fixed A, $\mu_1(A_{a_2})$ and $\mu_2(A_{a_1})$ can be regarded as functions of a_2 and a_1, respectively. The lemma below asserts that under some restrictions they are even measurable.

9.10 Lemma. *Let μ_1 and μ_2 be σ-finite measures on Σ_1 and Σ_2, respectively. Then for a fixed set $A \in \Sigma_1 \otimes \Sigma_2$, the functions*

$$f_A(a_1) = \mu_2(A_{a_1}) \text{ and } g_A(a_2) = \mu_1(A_{a_2})$$

are Σ_1-$\overline{\mathcal{B}}_+$- and Σ_2-$\overline{\mathcal{B}}_+$-measurable,respectively.

Proof.

(i) We first prove this statement under the assumption that μ_2 is finite. Let $\mathfrak{A} = \{A \in \Sigma_1 \otimes \Sigma_2 : f_A \text{ is } \Sigma_1\text{-}\mathcal{B}_+\text{-measurable}\}$. We show that $\mathfrak{A} = \Sigma_1 \otimes \Sigma_2$.

$a)$ For $\Omega = \Omega_1 \times \Omega_2$ and $f_\Omega(a_1) = \mu_2(\Omega_2)$ (is a nonnegative real number), we have that f_Ω is $\{\Omega, \emptyset\}$-$\mathcal{B}_+$-measurable, thus also Σ_1-$\mathcal{B}_+$-measurable, and hence $\Omega \in \Sigma$. Observe that the finiteness of μ_2 is essential, because otherwise, we can only arrive at the weaker result that f_A is Σ_1-$\overline{\mathcal{B}}_+$-measurable.

$b)$ Let $A \in \mathfrak{A}$. Then f_A is measurable and

$$A^c \in \Sigma_1 \otimes \Sigma_2.$$

Then, due to Problem 9.2,

$$f_{A^c}(a_1) = \mu_2((A^c)_{a_1}) = \mu_2((A_{a_1})^c)$$

$$= \mu_2(\Omega_2) \; - \mu_2(A_{a_1}) = f_\Omega(a_1) - f_A(a_1)$$

is measurable as the difference of two measurable functions. This implies that $A^c \in \mathfrak{A}$.

$c)$ Let $\{A_n\}$ be a sequence of disjoint sets from $\mathfrak{A}$. Then, again by Problem 9.2,

$$f_{\Sigma_{n=1}^\infty A_n}(a_1) = \mu_2((\textstyle\sum_{n=1}^\infty A_n)_{a_1}) = \mu_2(\textstyle\sum_{n=1}^\infty (A_n)_{a_1})$$

$$= \textstyle\sum_{n=1}^\infty \mu_2((A_n)_{a_1}) = \sum_{n=1}^\infty f_{A_n}(a_1) = \sup\{\textstyle\sum_{i=1}^n f_{A_i}(a_1)\}$$

is clearly measurable. This implies that $\sum_{n=1}^\infty A_n \in \mathfrak{A}$ and thus $\mathfrak{A}$ is a Dynkin system.

$d)$ Now let $A = A_1 \times A_2$ be a rectangle. Then by equation (9.9),

$$f_A(a_1) = \mu_2((A_1 \times A_2)_{a_1}) = \mu_2(A_2), \text{ if } a_1 \in A_1 \text{ and}$$

$$f_A(a_1) = 0, \text{ if } a_1 \notin A_1;$$

and therefore,

$$f_A(a_1) = \mu_2(A_2)\mathbf{1}_{A_1}(a_1) \tag{9.10}$$

is Σ_1-$\mathcal{B}_+$-measurable as the product of a constant function $\mu_2(A_2)$ and an indicator function of a measurable set $A_1 \in \Sigma_1$. Thus, $\mathfrak{A}$ contains the set $\Sigma_1 \times \Sigma_2$ of all measurable rectangles which is $\cap$-stable. On the other hand, $\Sigma_1 \otimes \Sigma_2$ is a Dynkin system generated by an $\cap$-stable generator, which implies $\Sigma_1 \otimes \Sigma_2 \subseteq \mathfrak{A}$ (as $\mathfrak{A}$ is a Dynkin system just containing $\Sigma_1 \times \Sigma_2$). However, by the definition of $\mathfrak{A}$, the family $\mathfrak{A} \subseteq \Sigma_1 \otimes \Sigma_2$. Therefore, $\mathfrak{A}$ is a σ-algebra and

$$\mathfrak{A} = \Sigma_1 \otimes \Sigma_2.$$

(ii) Now we assume that μ_2 is σ-finite. Then there exists a sequence

$$\{\Omega_2^k\} \subseteq \Sigma_2$$

such that

$$\Omega_2^k \uparrow \Omega_2 \quad \text{and} \quad \mu_2(\Omega_2^k)$$

is finite and we can apply the above result to the trace σ-algebras $\Sigma_1 \cap \Omega_2^k$, and state that f_A^k is

$$\Sigma_1 \cap \Omega_2^k\text{-}\mathcal{B}_+\text{-measurable},$$

where

$$f_A^k(a_1) = \mu_2(A_{a_1} \cap \Omega_2^k) \quad \text{and} \quad A \in \Sigma_1 \otimes \Sigma_2.$$

Consequently, by continuity from below for μ_2, the function

$$\mu_2(A_{a_1}) = f_A = \sup\{f_A^k\} = \lim_{k \to \infty} \mu_2(A_{a_1} \cap \Omega_2^k)$$

is Σ_1-$\overline{\mathcal{B}}_+$-measurable. □

Now, we come to a major result in this section about the formation of a product measure.

9.11 Theorem. *Let $(\Omega_i, \Sigma_i, \mu_i)$ be two measure spaces and let μ_i be σ-finite measures. Then there exists a measure μ on $\Sigma_1 \otimes \Sigma_2$*

such that, for each $A \in \Sigma_1 \otimes \Sigma_2$, $\mu(A)$ *is determined by the formula*

$$\mu(A) = \int \mu_2(A_{a_1}) \mu_1(da_1) = \int \mu_1(A_{a_2}) \mu(da_2). \tag{9.11}$$

Specifically, for $A = A_1 \times A_2$,

$$\mu(A_1 \times A_2) = \mu_1(A_1) \mu_2(A_2). \tag{9.11a}$$

Proof. Let $f_A(a_1) = \mu_2(A_{a_1})$. Then, by Lemma 9.10, f_A is a nonnegative extended-valued Σ_1-$\overline{\mathcal{B}}_+$-measurable function. Therefore, the integral

$$A \mapsto \int f_A(a_1) \mu_1(da_1) \tag{9.11b}$$

is defined for any $\Sigma_\otimes$-measurable set A and it is a nonnegative set function denotes by $\mu(A)$. We will prove that μ is a measure on $\Sigma_\otimes$. Clearly $\mu(\emptyset) = 0$. Using Problem 9.3 we can show that

$$f_{\sum_{n=1}^{\infty} A_n} = \sum_{n=1}^{\infty} f_{A_n}$$

for $\{A_n\}$ as a sequence of measurable pairwise disjoint sets of $\Sigma_\otimes$. Applying this equation to (9.11b), we have, by Beppo Levi's Corollary 2.2,

$$\begin{aligned} \mu\left(\sum_{n=1}^{\infty} A_n\right) &= \int \sum_{n=1}^{\infty} f_{A_n}(a_1) \mu_1(da_1) \\ &= \sum_{n=1}^{\infty} \int f_{A_n} d\mu_1 = \sum_{n=1}^{\infty} \mu(A_n). \end{aligned}$$

In particular, for $A = A_1 \times A_2$,

$$\mu(A_1 \times A_2) = \int f_A(a_1) \mu_1(da_1)$$

by (9.10)

$$= \int \mu_2(A_2) 1_{A_1}(a_1) \mu_1(da_1) = \mu_1(A_1) \mu_2(A_2). \qquad \square$$

The measure μ on $\Sigma_1 \otimes \Sigma_2$ is called a *product measure* and it is denoted by $\mu_1 \otimes \mu_2$. Theorem 9.11 is readily extendible to the product measure $\bigotimes_{i=1}^{n} \mu_i$ on finitely many measure spaces.

The uniqueness of the product measure is subject to the following.

9.12 Theorem. *Let $(\Omega_i, \Sigma_i, \mu_i)$ be measure spaces such that $\mathcal{G}_i$ are $\cap$-stable generators of Σ_i, $i = 1, \ldots, n$. If μ_i's are σ-finite, then the product measure $\bigotimes_{i=1}^{n} \mu_i$ on $\bigotimes_{i=1}^{n} \Sigma_i$ is also σ-finite and unique.*

Proof. Denote $\mathcal{G} = \bigtimes_{i=1}^{n} \mathcal{G}_i$, which, according to Lemma 9.5, is a generator of $\Sigma_\otimes = \bigotimes_{i=1}^{n} \Sigma_i$. It is easily seen that $\mathcal{G}$ is $\cap$-stable. Indeed, using Problem 5.10 (a), Chapter 1,

$$\begin{aligned} & (G_1 \times \cdots \times G_n) \cap (H_1 \times \cdots \times H_n) \\ = \; & (G_1 \cap H_1) \times \cdots \times (G_n \cap H_n), \end{aligned}$$

in as much as each $\mathcal{G}_i$ is $\cap$-stable.

Because μ_i is σ-finite for each i, there are sequences

$$\widetilde{G}_{ik} \uparrow \Omega_i\,, \; i = 1, \ldots, n, \;\text{ and }\; k = 1, 2, \ldots,$$

such that $\mu_i(\widetilde{G}_{ik}) < \infty$. Obviously,

$$\Omega^{(k)} = \prod_{i=1}^{n} \widetilde{G}_{ik} \uparrow \Omega = \prod_{i=1}^{n} \Omega_i.$$

Furthermore,

$$\Big(\bigotimes_{i=1}^{n} \mu_i\Big)(\Omega^{(k)}) = \mu_1(\widetilde{G}_{1n}) \cdot \cdots \cdot \mu_n(\widetilde{G}_{nk}) < \infty.$$

Therefore, $\bigotimes_{i=1}^{n} \mu_i$ is σ-finite. Now we recall again that the product measure $\bigotimes_{i=1}^{n} \mu_i$ has the property

$$\left(\bigotimes_{i=1}^{n}\mu_i\right)\left(\prod_{i=1}^{n}R_i\right)=\mu_1(R_1)\cdots\mu_n(R_n)$$

for every measurable rectangle from $\Sigma_{\otimes}$. Specifically, this property holds for every rectangle selected from the generator $\mathcal{G}$. If there is another measure $\hat{\mu}$ on $\Sigma_{\otimes}$ with the same property on $\mathcal{G}$, in other words, $\bigotimes_{i=1}^{n}\mu_i=\hat{\mu}$ on $\mathcal{G}$, then, by Theorem 2.13, Chapter 5 (on the uniqueness of the extension of a measure), these two measures must coincide on $\Sigma_{\otimes}$. This proves the uniqueness of the product measure. Observe that another measure on $\Sigma_{\otimes}$ may exist. However, it cannot be a product measure. □

9.13 Remark. Summarizing all results and formalism from Remark 9.4 throughout Theorem 9.12 we pass through the following major steps.

(i) Given a finite family of measure spaces $\{(\Omega,\Sigma_i,\mu_i): i=1,\ldots,n\}$ we defined the set $\bigtimes_{i=1}^{n}\Sigma_i$ of all measurable rectangles which turned out to be a semi-ring on $\Omega=\prod_{i=1}^{n}\Omega_i$.

(ii) We defined on $\bigtimes_{i=1}^{n}\Sigma_i$ an elementary content μ^0 by

$$\mu^0\big(R^{(n)}\big)=\mu_1(A_1)\cdots\mu_n(A_n), \tag{9.13}$$

which was extendible to a content $\widetilde{\mu}$ on the algebra $\mathcal{A}\Big(\bigtimes_{i=1}^{n}\Sigma_i\Big)$ of all n-dimensional figures F representable as a finite union of disjoint rectangles: $F=\sum_{k=1}^{m}R_k^{(n)}$ so that each such figure could be measured by $\widetilde{\mu}$ as $\widetilde{\mu}(F)=\sum_{k=1}^{m}\mu^0(R_k^{(n)})$.

(iii) An extension of $\widetilde{\mu}$ from $\mathcal{A}\Big(\bigtimes_{i=1}^{n}\Sigma_i\Big)$ to $\Sigma_{\otimes}$ was rendered explicitly via integration. More precisely, the common transition from the elementary content on the semi-ring to the content and then a measure was cut short by defining the measure

$$\mu(A)=\int\mu_2(\sigma_{a_1}(A))\mu_1(da_1)=\int\mu_1(\sigma_{a_2}(A))\mu(da_2)$$

on $\Sigma_{\otimes}$ in (9.11) that agreed with the content μ^0 (9.13) when μ was applied to rectangles. To avoid bulky notation, the results were established for two factor spaces.

(iv) According to Theorem 9.12, measure μ turned out to be σ-finite, which guaranteed the uniqueness of such a product measure μ, also in notation $\bigotimes_{i=1}^{n} \mu_i$. Of course, we should have mentioned that μ_i's were assumed to be σ-finite to begin with.

(v) In conclusion, we note that although some elements of the Carathéodory extension were present in the formation of the product measure, it was not exactly the one we exhibited in Section 2, Chapter 5. For example, in one case of the product measure on Borel σ-algebras in Euclidean spaces, the Borel - Lebesgue measure λ^n on $\mathcal{B}^n$ and the product measure $\bigotimes_{i=1}^{n} \lambda_i$ on $\mathcal{B}_{\otimes}$ must be identical (Problem 9.8). However, the Lebesgue measure λ_0^{n*} on the Lebesgue σ-algebra $\mathcal{L}^*(\mathbb{R}^n)$ and the product Lebesgue measure $\bigotimes_{i=1}^{n} \lambda_i^*$ do not coincide, as we learn from the example below. This was one of the reasons why we cared to distinguish the Borel - Lebesgue and Lebesgue measures. It would be fruitful to further investigate the differences between topologically induced σ-algebras and Borel or Borel-like measures compared to those stemming from Carathéodory extensions on finite product spaces. However, this is beyond the scope of this book. □

9.14 Example. Recall that the Lebesgue measure λ_0^* is complete on Lebesgue σ-algebra $\mathcal{L}^*$. Consider Lebesgue measure λ_0^* on $\mathcal{L}^*(\mathbb{R})$. We show that $\lambda_0^* \otimes \lambda_0^*$ is not complete. Let Q be any subset of $\mathbb{R}$ that does not belong to $\mathcal{L}^*(\mathbb{R})$. Then, $\{0\} \times Q \notin \mathcal{L}^* \otimes \mathcal{L}^*$ or else, by Problem 9.4, Q would be an element of $\mathcal{L}^*$. On the other hand, $\{0\} \times Q$ is a (proper) subset of $\{0\} \times \mathbb{R}$. The latter set, by Problem 3.1, Chapter 5, is a measurable Borel null set. Hence, $\{0\} \times Q$ is a negligible non measurable set in $\mathcal{L}^* \otimes \mathcal{L}^*$, although it is measurable in the Lebesgue space $(\mathbb{R}^2, \mathcal{L}^*(\mathbb{R}^2), \lambda_0^{2*})$, known to be complete. Hence, the space $(\mathbb{R}^2, \mathcal{L}^*(\mathbb{R}^2), \lambda_0^{2*})$ does not coincide with the product measure space $(\mathbb{R}^2, \mathcal{L}^* \otimes \mathcal{L}^*, \lambda_0^* \otimes \lambda_0^*)$. Obviously, the same discrepancy takes place in the n-dimensional Euclidean spaces

$(\mathbb{R}^n, \mathcal{L}^*(\mathbb{R}^n), \lambda_0^{*n})$ and $\left(\mathbb{R}^n, \overset{n}{\underset{i=1}{\otimes}} \mathcal{L}_i^*(\mathbb{R}), \overset{n}{\underset{i=1}{\otimes}} \lambda_{0i}^*\right)$, in contrast with its Borel - Lebesgue counterpart. □

After we define the product measure on a measurable product space we turn to multiple integration of measurable functions on product spaces with respect to the product measure.

9.15 Definition. Let $f : \Omega_1 \times \Omega_2 \to \Omega_3$ be a function, $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$ and let π_1, π_2 and π_3 be the corresponding projection maps in Ω. Let

$$\pi_{23} : \Omega \to \Omega_2 \times \Omega_3$$

be the projection map defined as $\pi_{23}(\omega_1, \omega_2, \omega_3) = (\omega_2, \omega_3)$. For a distinction, we now use a different character for the section operator applied to functions from that we used for sets. Denote

$$\varphi_{a_1} \circ f = f_{a_1} = \pi_{23*}(\pi_1^*(a_1) \cap f)$$

and call it the a_1-*section of* f. In other words,

$$\varphi_{a_1} \circ f = \{(\omega_2, f(a_1, \omega_2)) : \omega_2 \in \Omega_2\}$$

or also in the form

$$f_{a_1} = \{(\omega_2, \omega_3) \in \Omega_2 \times \Omega_3 \colon (a_1, \omega_2, \omega_3) \in f\}.$$

The a_2-section is defined similarly. □

9.16 Proposition. *Let* $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$ *be a measurable product space*, (Ω_3, Σ_3) *be a measurable space*, *and let* $f : \Omega_1 \times \Omega_2 \to \Omega_3$ *be a* $\Sigma_1 \otimes \Sigma_2$-$\Sigma_3$*-measurable function. Then the section* f_{a_1} *is* Σ_2-Σ_3*-measurable and* f_{a_2} *is* Σ_1-Σ_3*-measurable.*

(See Problem 9.13.)

The theorems below resemble some of the original versions of Fubini's theorem. Theorem 9.17 (which by many is referred to as *Tonelli's theorem*) states in essence that if the integral of a non-negative extended-valued $\Sigma_1 \otimes \Sigma_2$-measurable function is finite,

then it can be calculated by using iterated integrals of formula (9.17) below. To check whether an arbitrary $\Sigma_1 \otimes \Sigma_2$-measurable function f is integrable one can also apply Theorem 9.17 directly to $|f|$ by using formula (9.17).

9.17 Theorem (Tonelli). *Let* $(\Omega_i, \Sigma_i, \mu_i),\ i = 1, 2,$ *be* σ*-finite measure spaces and let* $f \in \mathcal{C}_+^{-1}(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$. *Then the functions*

$$a_1 \mapsto \int f_{a_1}(\omega_2)\, d\mu_2(\omega_2) \ \text{ and } \ a_2 \mapsto \int f_{a_2}(\omega_1) d\mu_1(\omega_1)$$

are Σ_1*- and* Σ_2*-measurable, respectively, and*

$$\begin{aligned} \int f d\mu_1 \otimes \mu_2 &= \int (\int f_{a_1}(\omega_2) d\mu_2(\omega_2)) d\mu_1(a_1) \\ &= \int (\int f_{a_2}(\omega_1) d\mu_1(\omega_1)) d\mu_2(a_2). \end{aligned} \tag{9.17}$$

Proof.

(i) As usual, we begin with nonnegative simple functions. Let

$$s = \textstyle\sum_{i=1}^n \alpha_i \mathbf{1}_{A_i},\ A_i \in \Sigma_1 \otimes \Sigma_2.$$

By Problems 9.5 and 9.14,

$$s_{a_1}(\omega_2) = \textstyle\sum_{i=1}^n \alpha_i (\mathbf{1}_{A_i})_{a_1}(\omega_2) = \sum_{i=1}^n \alpha_i \mathbf{1}_{(A_i)_{a_1}}(\omega_2),$$

and the latter is measurable and

$$\int s_{a_1}(\omega_2) d\mu_2(\omega_2) = \textstyle\sum_{i=1}^n \alpha_i \mu_2((A_i)_{a_1}),$$

which is Σ_1-$\mathcal{B}_+$-measurable by Lemma 9.9. Hence we may integrate $\int s_{a_1}(\omega_2) d\mu_2(\omega_2)$ with respect to measure μ_1:

$$\begin{aligned} &\int \Big(\int s_{a_1}(\omega_2) d\mu_2(\omega_2) \Big)\, d\mu_1(a_1) \\ &= \textstyle\sum_{i=1}^n \alpha_i \int \mu_2((A_i)_{a_1})\, d\mu_1(a_1). \end{aligned} \tag{9.17a}$$

By Theorems 9.11 and 9.12, equation (9.17a) reduces to

$$\sum_{i=1}^{n} \alpha_i \mu_1 \otimes \mu_2(A_i) = \int s d\mu_1 \otimes \mu_2 \,. \tag{9.17b}$$

(*ii*) Let $f \in \mathcal{C}_+^{-1}(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$. Then there is a sequence $\{s_n\}$ such that $\{s_n\}\uparrow \subseteq \Psi_+(\Omega_1 \times \Omega_2, \mathcal{C}^{-1})$ and that $f = \sup\{s_n\}$. Thus

$$\int f d\mu_1 \otimes \mu_2 = \sup\Big\{\int s_n \, d\mu_1 \otimes \mu_2\Big\}$$

(by (9.17a) and (9.17b))

$$= \sup \int (\int (s_n(\omega_2))_{a_1} d\mu_2(\omega_2)) d\mu_1(a_1)$$

(by the monotone convergence theorem)

$$= \int (\sup \int (s_n(\omega_2))_{a_1} \, d\mu_2(\omega_2)) d\mu_1(a_1).$$

By Problem 9.12, finally,

$$\int f d\mu_1 \otimes \mu_2 \; = \int (\int f_{a_1}(\omega_2) d\mu_2(\omega_2)) d\mu_1(a_1).$$

Now it is obvious how to complete the proof of formula (9.17). □

9.18 Theorem (Fubini). *Let* $(\Omega_i, \Sigma_i, \mu_i)$, $i = 1, 2$, *be* σ*-finite measure spaces and let* $f \in L^1(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2; \overline{\mathbb{R}})$. *Then the functions*

$$\omega_2 \;\mapsto\; f_{a_1}(\omega_2) \quad \textit{and} \quad \omega_1 \;\mapsto\; f_{a_2}(\omega_1)$$

are μ_2*-integrable* μ_1-a.e., *and* μ_1*-integrable* μ_2-a.e., *respectively*; *the functions*

$$a_1 \;\mapsto\; \int f_{a_1}(\omega_2) \, d\mu_2(\omega_2) \;\text{ and }\; a_2 \;\mapsto\; \int f_{a_2}(\omega_1) d\mu_1(\omega_1)$$

are μ_1-a.e. *and* μ_2-a.e. *defined*, *and they are* μ_1*- and* μ_2*- integrable*, *respectively*, *and the formula*

$$\int f d\mu_1 \otimes \mu_2 = \int (\int f_{a_1}(\omega_2) d\mu_2(\omega_2)) d\mu_1(a_1)$$

$$= \int (\int f_{a_2}(\omega_1) d\mu_1(\omega_1)) d\mu_2(a_2) \tag{9.18}$$

holds.

Proof. Because by the assumption of the theorem,

$$f \in L^1(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2; \overline{\mathbb{R}}),$$

we have by Problem 9.13 and then by Tonelli's Theorem 9.17 applied to $|f|$ that

$$\int |f| d\mu_1 \otimes \mu_2 = \int (\int |f_{a_1}(\omega_2)| d\mu_2(\omega_2)) d\mu_1(a_1)$$

$$= \int (\int |f_{a_2}(\omega_1)| d\mu_1(\omega_1)) d\mu_2(a_2) < \infty.$$

By Proposition 1.21, the functions $a_1 \mapsto \int |f_{a_1}(\omega_2)| d\mu_2(\omega_2)$ and $a_2 \mapsto \int |f_{a_2}(\omega_1)| d\mu_1(\omega_1)$ are μ_1- and μ_2-a.e. finite, respectively. In other words, for almost all $a_1 \in \Omega_1$ and $a_2 \in \Omega_2$, $|f_{a_1}|$ and $|f_{a_2}|$ are μ_2- and μ_1-integrable, respectively. It is easy to verify that the same applies to the pairs of functions $(f_{a_1})^+$, $(f_{a_1})^-$ and $(f_{a_2})^+$, $(f_{a_2})^-$. By Theorem 9.17, the pairs of functions

$$a_1 \mapsto \int (f_{a_1})^+(\omega_2) d\mu_2(\omega_2),\ a_1 \mapsto \int (f_{a_1})^-(\omega_2) d\mu_2(\omega_2)$$

and

$$a_2 \mapsto \int (f_{a_2}(\omega_1))^+ d\mu_1(\omega_1),\ a_2 \mapsto \int (f_{a_2}(\omega_1))^- d\mu_1(\omega_1)$$

are Σ_1- and Σ_2-measurable, and they are μ_1-a.e. and μ_2-a.e. defined, respectively. Finally, applying formula (9.17) to f^+, $(f_{a_1})^+$, and $(f_{a_2})^+$ and then to f^-, $(f_{a_1})^-$, and $(f_{a_2})^-$, we arrive at formula (9.17). □

9.19 Remarks.

(i) Fubini's theorem generalizes Theorem 9.11 because for $f = \mathbf{1}_A$ and $A \in \Sigma_1 \otimes \Sigma_2$, the result of Theorem 9.11 immediately

follows. This is why Theorem 9.11 is also referred to as Fubini's theorem.

(ii) Observe that Tonelli's and Fubini's theorems differ not only in that they are applied to nonnegative and arbitrary measurable functions, respectively, but also that functions on a product space belong to L^1-space is a conclusion in Tonelli's Theorem whereas it is a hypothesis in Fubini's theorem.

(iii) The above results (including Fubini's theorem) can naturally be extended from two spaces to any finitely many spaces. This involves a relatively cumbersome notation and is not discussed further except by way of examples. □

9.20 Definition. Let $(\Omega_i, \Sigma_i, \mu_i),\ \ i = 1, \ldots, n,$ be measure spaces with σ-finite measures. Then the triple

$$\Big(\prod_{i=1}^{n} \Omega_i\,,\ \overset{n}{\underset{i\,=\,1}{\otimes}} \Sigma_i,\ \overset{n}{\underset{i\,=\,1}{\otimes}} \mu_i\Big)$$

is called the *product* (*measure*) *space*. □

9.21 Examples.

(i) Let $C_n(x_0\,, r)$ be a closed ball in $(\mathbb{R}^n, \tau_n)$. Denote $V_n = \lambda^n(C_n(\theta, 1))$. We wish to show that $\lambda^n(C_n(x_0,\, r)) = V_n \cdot r^n,$ where

$$V_{2k} = \tfrac{1}{k!}\,\pi^k\,,\ k = 1, 2, \ldots \tag{9.21}$$

and

$$V_{2k-1} = \tfrac{2^k}{1\cdot 3\cdots(2k-1)}\,\pi^{k-1}\,,\ k = 1, 2, \ldots\ . \tag{9.21a}$$

A closed ball is clearly a Borel set. Let $L_n(r,\, x_0)$: $\mathbb{R}^n \to \mathbb{R}^n$ be the bijective map,

$$L(r, x_0)(\,x) = rx + x_0,$$

where $x,\, x_0 \in \mathbb{R}^n$ and $r > 0$. Then, $C_n(\theta, 1) = L^*(r, x_0)(C_n(r,\, x_0))$. Observe that

$$L^*(r, x_0) = E(\tfrac{1}{r}) \circ M(-x_0),$$

where E means expansion with factor $\frac{1}{r}$ and M stands for the parallel motion (i.e. translation, here with the shift $-x_0$). Now,

$$C_n(\theta, 1) = \{x \in \mathbb{R}^n \colon \| x \|^2 \leq 1\}$$

$$= \{(x_1, \ldots, x_n) \colon \sum_{i=1}^{n} x_i^2 \leq 1\}$$

$$= \{(x_1, \ldots, x_n) \colon \sum_{i=3}^{n} x_i^2 \leq 1 - x_1^2 - x_2^2\,, \quad \text{with } (x_1, x_2) \in C_2(\theta, 1)\}$$

$$= C_{n-2}\big(\theta, \sqrt{1 - x_1^2 - x_2^2}\big)$$

$$= L^*_{n-2}((1 - x_1^2 - x_2^2)^{-1/2}, \theta)(C_{n-2}(\theta, 1))\,,$$

where $(x_1, x_2) \in C_2(\theta, 1)$. This yields

$$\mathbf{1}_{C_n(\theta,1)}(x_1, \ldots, x_n)$$

$$= \mathbf{1}_{C_{n-2}(\theta, \sqrt{1-x_1^2-x_2^2})}(x_3, \ldots, x_n) \cdot \mathbf{1}_{C_2(\theta,1)}(x_1, x_2)$$

$$= \mathbf{1}_{L^*_{n-2}((1-x_1^2-x_2^2)^{1/2}, \theta)(C_{n-2}(\theta,1))}(x_3, \ldots, x_n)$$

$$\times \mathbf{1}_{C_2(\theta,1)}(x_1, x_2)\,. \tag{9.21b}$$

Furthermore,

$$V_n = \int\limits_{C_n(\theta,1)} \lambda^n(dx) = \int \mathbf{1}_{C_n(\theta,1)}(x)(\lambda^{n-2} \otimes \lambda^2)(dx)$$

[by Fubini's theorem and by (9.21b)]

$$= \int \mathbf{1}_{C_2(\theta,1)}(x_1, x_2)$$

$$\times \int \mathbf{1}_{L^*_{n-2}((1-x_1^2-x_2^2)^{1/2},\theta)(C_{n-2}(\theta,1))} d\lambda^{n-2} d\lambda^2(x_1, x_2) \, .$$

The interior (second) integral is, due to Proposition 4.3, part 1), Chapter 5, and the above observation, equal to

$$\int \mathbf{1}_{C_{n-2}(\theta,1)} \circ E((1 - x_1^2 - x_2^2)^{-1/2})(x_3, \ldots, x_n) d\lambda^{n-2}(x_3, \ldots, x_n) \, .$$

By Proposition 4.3, part 2), Chapter 5 and by Theorem 4.1, the last integral equals

$$(1 - x_1^2 - x_2^2)^{(n-2)/2} \, V_{n-2} \, .$$

Therefore,

$$V_n = V_{n-2} \int (1 - x_1^2 - x_2^2)^{(n-2)/2} \, \mathbf{1}_{C_2(\theta,1)}(x_1, x_2) d\lambda^2(x_1, x_2)$$

(by Fubini's theorem)

$$= \, V_{n-2} \int\int (1 - x_1^2 - x_2^2)^{(n-2)/2} \, \mathbf{1}_{C_2(0,1)}(x_1, x_2) d\lambda(x_1) d\lambda(x_2) \, .$$

This is a Lebesgue integral of a continuous function on the unit ball and it can be reduced to a Riemann integral by using conventional techniques for Riemann integrals. For example, the double integral above is then

$$\int\limits_{\rho \in [a,x]} \int\limits_{\theta \in [0,\pi]} \rho (1 - \rho^2)^{(n-2)/2} \, d\theta d\rho = \tfrac{2\pi}{n} \, , \tag{9.21c}$$

and thus

$$V_n = V_{n-2} \, \tfrac{2\pi}{n}, \; n = 2, 3, \ldots \, . \tag{9.21d}$$

Let $V_0 = 1$. Then, $V_1 = 2$ (as the Lebesgue measure of the interval $[-1, 1]$). By (9.21c), $V_2 = \pi$ (which agrees with the definition of V_0).

$$V_3 = \, V_1 \, \tfrac{2\pi}{3} = \tfrac{2^2}{1 \cdot 3} \, \pi^{2-1} \, , \text{and } \; V_4 = \, \tfrac{\pi^2}{2!} .$$

The validity of formulas (9.21) and (9.21a) is then easily shown by induction and the use of (9.21d).

(*ii*) We show that Fubini's theorem need not hold when at least one of the measures, μ_1 or μ_2, is not σ-finite. Let $(\Omega_i, \Sigma_i) = ([0,1], \mathcal{B}([0,1])),\ i = 1,2,\ \mu_1 = \mathrm{Res}_{[0,1]}\lambda$, and $\mu_2(A) = |A|$, if A is finite and $\mu_2(A) = \infty$, if A is infinite, where $A \in \Sigma_2$. Denote

$$D = \{(x_1, x_2) \in [0,1]^2 \colon x_1 = x_2\},$$

the diagonal of the square (see Figure 9.2 below). We show that $D \in \Sigma_1 \otimes \Sigma_2 = \mathcal{B}^2([0,1]^2)$. Let

$$S_{nj} = \left[\tfrac{j-1}{n},\ \tfrac{j}{n}\right],\ j = 1, \dots, n,$$

and

$$A_n = \bigcup_{k=1}^{n} (S_{nk} \times S_{nk}).$$

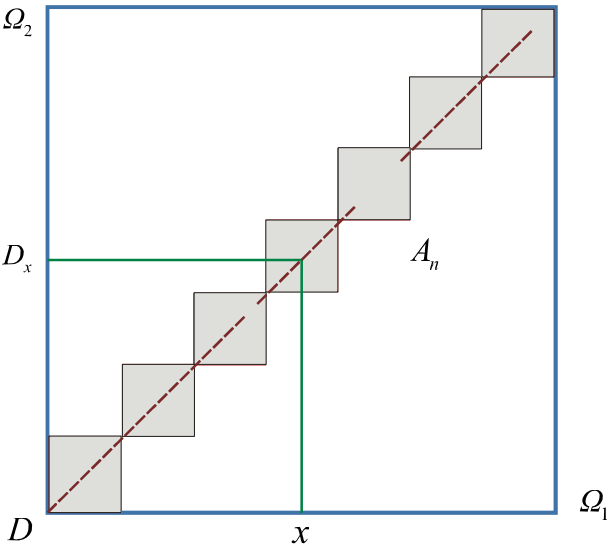


Figure 9.2

Then $D \in \Sigma_1 \otimes \Sigma_2$ for $D = \bigcap_{n=1}^{\infty} A_n$. Now we find

$$\int_{[0,1]} \mu_2(D_x)\lambda(dx) = \lambda([0,1]) = 1 \text{ (because } \mu_2(D_x) = 1).$$

On the other hand,

$$\int_{[0,1]} \lambda(D_y)\mu_2(dy) = 0.$$

So as we see, Fubini's theorem or more precisely, the second equation in (9.11) of Theorem 9.11, does not hold.

(iii) Let $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \gamma)$ be the counting measure space introduced in Example 1.2$(viii)$, Chapter 5, for more general measure spaces. We consider a sequence $\{s_n\}$ of nonnegative simple functions on $\mathbb{N}$ as

$$s_n = \textstyle\sum_{i=0}^{n} a_i \mathbf{1}_{\{i\}},\ n = 0, 1, \ldots,$$

where $\{a_k\}$ is a sequence of nonnegative reals, so that

$$g = \sup\{s_n\} = \textstyle\sum_{i=0}^{\infty} a_i \mathbf{1}_{\{i\}}.$$

Hence the integral of g turns to the series:

$$\textstyle\int g d\gamma = \sup \int s_n d\gamma = \sum_{i=0}^{\infty} a_i. \tag{9.21e}$$

This is readily extendible to a series with real-valued terms. In other words, the integral of a sequence $\{a_n\} \subseteq \mathbb{R}$ with respect to the counting measure γ is represented by the series in (9.21e).

Let $\{f_n\}$ be a sequence of nonnegative functions of $\mathcal{C}_+^{-1}(\Omega, \Sigma)$ and let μ be a σ-finite measure on Σ. The above counting measure γ is σ-finite, therefore the function f [where $f(n, \omega) = f_n(\omega)$] obviously meets the conditions of Tonelli's Theorem 9.17: $f \in \mathcal{C}_+^{-1}(\mathbb{N} \times \Omega, \mathcal{P}(\mathbb{N}) \otimes \Sigma)$. Consequently, the sections

$$n \mapsto \textstyle\int f_n(\omega) d\mu(\omega) \quad \text{and} \quad a \mapsto \int f(\,\cdot\,, a) d\gamma$$

are $\mathcal{P}(\mathbb{N})$- and Σ-measurable, respectively, and

$$\int f d\gamma \otimes \mu = \sum_{n=0}^{\infty} \int f_n d\mu = \int \sum_{n=0}^{\infty} f_n d\mu. \tag{9.21f}$$

Although (9.21f) is a nice illustration of Tonelli's theorem, it is a slightly weaker alternative to Beppo Levi's Corollary 2.2, since the latter does not require μ to be σ-finite.

Now, let $\{f_n\} \in \mathcal{C}^{-1}(\Omega, \Sigma)$. To use an analogue of Fubini's theorem, we need to make sure that the function f belongs to $L^1(\mathbb{N} \times \Omega, \mathcal{P}(\mathbb{N}) \otimes \Sigma, \gamma \otimes \mu; \overline{\mathbb{R}}_+)$, or, alternatively, apply the above procedure initially to the sequence $\{|f_n|\}$ instead. Then, from (9.21f) we can get

$$\begin{aligned} \int |f| d\gamma \otimes \mu &= \sum_{n=0}^{\infty} \int |f_n| d\mu \\ &= \int \sum_{n=0}^{\infty} |f_n| d\mu. \end{aligned} \tag{9.21g}$$

Should now, $\int |f| d\gamma \otimes \mu$ or their equivalents, $\sum_{n=0}^{\infty} \int |f_n| d\mu$ or $\int \sum_{n=0}^{\infty} |f_n| d\mu$, be finite, it would yield that

$$f \in L^1(\mathbb{N} \times \Omega, \mathcal{P}(\mathbb{N}) \otimes \Sigma, \gamma \otimes \mu; \overline{\mathbb{R}})$$

and therefore, Fubini's formula (9.21f) would hold true, now for an arbitrary sequence of measurable functions $\{f_n\}$. Notice that, because

$$\int |f| d\gamma \otimes \mu < \infty$$

is a necessary condition for $f \in L^1(\mathbb{N} \times \Omega, \mathcal{P}(\mathbb{N}) \otimes \Sigma, \gamma \otimes \mu; \overline{\mathbb{R}})$, it automatically implies that

$$\sum_{n=0}^{\infty} \int |f_n| d\mu < \infty \tag{9.21h}$$

or

$$\int \sum_{n=0}^{\infty} |f_n| d\mu < \infty$$

would be alternative necessary conditions for f to belong to the space $L^1(\mathbb{N}\times\Omega, \mathcal{P}(\mathbb{N})\otimes\Sigma, \gamma\otimes\mu; \overline{\mathbb{R}})$ (although the latter is in no way a necessary condition for Fubini's theorem). In some applications, this version of Fubini's theorem can challenge the generalized monotone convergence Theorem 2.4 and Lebesgue's dominated convergence Theorem 2.6.

(iv) As an illustration to the last application of Fubini's theorem, consider a r.v. X on a probability space $(\Omega, \Sigma, \mathbb{P})$. The function $m(\theta) \mapsto \mathbb{E}[e^{\theta X}]$ (normally, complex-valued) is known to be the *moment generating function* of X. If we formally expand $e^{\theta X}$ in the Maclaurin series,

$$e^{\theta X} = \sum_{n=0}^{\infty} \tfrac{\theta^n}{n!} X^n$$

assuming $m(\theta)$ is analytic at zero, we have with

$$m(\theta) = \mathbb{E}\big[\sum_{n=0}^{\infty} \tfrac{\theta^n}{n!} X^n\big]$$

a precedent for the application of Fubini's theorem discussed in Example (iii). Hence, we have to make sure that, in light of (9.21h), the series

$$\sum_{n=0}^{\infty} \mathbb{E}\Big[\tfrac{|\theta|^n}{n!}|X|\Big] = \sum_{n=0}^{\infty} \tfrac{|\theta|^n}{n!}\mathbb{E}[|X|^n]$$

is convergent in some vicinity of $\theta = 0$. [The latter holds if $m(\theta)$ is analytic at zero yielding that $\mathbb{E}[|X|^n] < \infty$ for all n.] Assuming that all absolute moments of X exist and the above series does converge, the application of Fubini's theorem (9.21f) yields that

$$m(\theta) = \sum_{n=0}^{\infty} \tfrac{\theta^n}{n!}\mathbb{E}[X^n],$$

as a Taylor series expansion of $m(\theta)$ in terms of all moments of the random variable X, and consequently that

$$\mathbb{E}[X^n] = m^{(n)}(0),\ n = 0, 1, \dots\ .$$

(v) Consider Borel - Lebesgue measure $\lambda^2 = \lambda \otimes \lambda$ on Borel σ-algebra $\mathcal{B}^2$. Let $A = \mathbb{Q} \times \mathbb{R}$. According to Problem 3.1, Chapter 5, A is a countable union of Borel-null sets. Thus,

$$\lambda^2(A) = \int \mathbf{1}_A d\lambda^2 = 0.$$

On the other hand, the section $(\mathbf{1}_A)_{a_1}$ is not λ-integrable for all $a_1 \in \mathbb{Q}$. This is, however, in agreement with Fubini's theorem that the function $(\mathbf{1}_A)_{a_1}$ is λ-integrable only for almost all $a_1 \in \mathbb{R}$.

(vi) Now we discuss yet another application often occurring in probability theory. Let μ_F and μ_G be two finite Borel - Lebesgue - Stieltjes measures induced by distribution functions $F, G \in \mathfrak{D}(\mathbb{R}, \mathcal{B})$ (see Remark 3.5(iii), Chapter 5). Recall that

$$\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} G(x) = 0.$$

From Problem 3.7, Chapter 5, given a compact interval $I = [a, b]$, we have that

$$\begin{aligned} \mu_F \otimes \mu_G(I^2) &= \mu_F(I)\mu_G(I) \\ &= [F(b) - F(a-)][G(b) - G(a-)]. \end{aligned} \tag{9.21i}$$

Let $T_u = \{(x, y) \in [a, b]^2 : y > x\}$ and $T_l = \{(x, y) \in [a, b]^2 : y \leq x\}$ be the upper and lower triangles of the square I^2, respectively. Now we calculate the measure of I^2 under $\mu_F \otimes \mu_G$ by using Theorem 9.11 in terms of Lebesgue - Stieltjes integrals:

$$\begin{aligned} \mu_F \otimes \mu_G([a, b]^2) &= \mu_F \otimes \mu_G(T_l) + \mu_F \otimes \mu_G(T_u) \\ &= \int_I \mu_G([a, x])\mu_F(dx) + \int_I \mu_F([a, y))\mu_G(dy) \\ &= \int_I [G(x) - G(a-)]\mu_F(dx) \\ &+ \int_I [F(x-) - F(a-)]\mu_G(dx). \end{aligned} \tag{9.21j}$$

Equating (9.21i) and (9.21j) we arrive at

$$\int_I G(x)\mu_F(dx) + \int_I F(x-)\mu_G(dx)$$

$$= F(b)G(b) - F(a-)G(a-). \tag{9.21k}$$

Interchanging the roles of F and G we have

$$\int_I G(x-)\mu_F(dx) + \int_I F(x)\mu_G(dx)$$

$$= F(b)G(b) - F(a-)G(a-). \tag{9.21l}$$

Hence, from (9.21k) and (9.21l) we establish the following integration by parts formula for Lebesgue - Stieltjes integrals.

$$F(b)G(b) - F(a-)G(a-)$$

$$= \int_I \tfrac{1}{2}\{F(x) + F(x-)\}\mu_G(dx)$$

$$+ \int_I \tfrac{1}{2}\{G(x) + G(x-)\}\mu_F(dx). \tag{9.21m}$$

□

9.22 Remark. In Chapter 5, sections 1, 4, and 5, we gave a rather extensive treatment of the notion of independence of families of events and r.v.'s. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and X_i be a (Ω, Σ)-(Ω_i, Σ_i) r.v., $i = 1, \ldots, n$. Recall that r.v.'s $X_1, \ldots, X_n$ are *independent* if the σ-algebras $\{\sigma(X_1), \ldots, \sigma(X_n)\}$ generated by these r.v.'s are independent. We have a new look into this notion.

First we denote $\overset{n}{\underset{i=1}{\otimes}} X_i = \{X_1, \ldots, X_n\} : \Omega \to \Omega_1 \times \cdots \times \Omega_n$ and call it the *product map*. Unlike Definition 4.12, Chapter 3, here the product map of functions is defined slightly differently:

$$\overset{n}{\underset{i=1}{\otimes}} X_i(\omega) = (X_1(\omega), \ldots, X_n(\omega)) = (\omega_1, \ldots, \omega_n).$$

Consequently, we specify

$$\left[\Omega, \prod_{i=1}^{n}\Omega_i, \overset{n}{\underset{i=1}{\otimes}} X_i\right]$$

as

$$\left(\overset{n}{\underset{i=1}{\otimes}} X_i\right)_*(A) = \{(X_1(\omega), \ldots, X_n(\omega)) : \omega \in A\}, A \subseteq \Omega$$

and

$$\left[\prod_{i=1}^{n}\Omega_i, \Omega, \left(\overset{n}{\underset{i=1}{\otimes}} X_i\right)^*\right]$$

as

$$\begin{aligned}&\left(\overset{n}{\underset{i=1}{\otimes}} X_i\right)^*(B)\\ &= \{X_1^*(\{\omega_1\}) \cap \cdots \cap X_n^*(\{\omega_n\}) : (\omega_1, \ldots, \omega_n) \in B\},\\ &B \subseteq \prod_{i=1}^{n}\Omega_i. \end{aligned} \tag{9.22}$$

In particular, if $B = B_1 \times \cdots \times B_n$,

$$\left(\overset{n}{\underset{i=1}{\otimes}} X_i\right)^*(B) = X_1^*(B_1) \cap \cdots \cap X_n^*(B_n). \tag{9.22a}$$

Indeed, if

$$A = \left(\overset{n}{\underset{i=1}{\otimes}} X_i\right)^*(B) = \left\{\omega \in \Omega : \overset{n}{\underset{i=1}{\otimes}} X_i(\omega) \in \prod_{i=1}^{n} B_i\right\},$$

we have

$$\begin{aligned}&\omega \in A \;\Leftrightarrow\; (X_1(\omega), \ldots, X_n(\omega)) \in \prod_{i=1}^{n} B_i\\ &\Leftrightarrow\; X_1(\omega) \in B_1, \ldots, X_n(\omega) \in B_n\end{aligned}$$

$$\Leftrightarrow \omega \in X_1^*(B_1) \cap \ldots \cap X_n^*(B_n).$$

It appears (Problem 9.22) that the product map $\bigotimes_{i=1}^{n} X_i$ is Σ-$\bigotimes_{i=1}^{n} \Sigma_i$-measurable. Therefore, by

$$\mathbb{P}_{\otimes X_i} = \mathbb{P}_{\bigotimes_{i=1}^{n} X_i} = \mathbb{P}\Big(\bigotimes_{i=1}^{n} X_i\Big)^*$$

we can define a probability measure on $\left(\prod_{i=1}^{n} \Omega_i\,, \bigotimes_{i=1}^{n} \Sigma_i\right)$ and call it the *joint distribution of r.v.'s* $X_1, \ldots, X_n$.

Denote $\mathbb{P}_{X_i} = \mathbb{P}X_i^*$ the distribution of the random variable X_i, $i = 1, \ldots, n$, which is a probability measure on Σ_i. Then, according to the construction we deal with in this section, we set the probability product space

$$\left(\prod_{i=1}^{n} \Omega_i, \bigotimes_{i=1}^{n} \Sigma_i, \bigotimes_{i=1}^{n} \mathbb{P}_{X_i}\right).$$

On the other hand, we already have another measure $\mathbb{P}_{\otimes X_i}$ on $\left(\prod_{i=1}^{n} \Omega_i, \bigotimes_{i=1}^{n} \Sigma_i\right)$ which, in general, need not be a product measure. The following statement sheds light on this dilemma. □

9.23 Theorem. *The joint probability distribution* $\mathbb{P}_{\otimes X_i}$ *is a product measure and equals* $\bigotimes_{i=1}^{n} \mathbb{P}_{X_i}$ *if and only if the r.v.'s* $X_1, \ldots, X_n$ *are independent.* □

Proof. (i) Suppose $\mathbb{P}_{\otimes X_i} = \bigotimes_{i=1}^{n} \mathbb{P}_{X_i}$. Then, as the product measure, $\bigotimes_{i=1}^{n} \mathbb{P}_{X_i}$ has the property that is factorized at all measurable rectangles, namely, given sets $A_1 \in \Sigma_1, \ldots, A_n \in \Sigma_n$,

$$\bigotimes_{i=1}^{n} \mathbb{P}_{X_i}(A_1 \times \cdots \times A_n) = \mathbb{P}_{X_1}(A_1) \cdots \mathbb{P}_{X_n}(A_n), \tag{9.23}$$

which by Problems 4.11 and 4.12 means that $X_1, \ldots, X_n$ are independent. Indeed, by the assumption, the left-hand side of (9.23) is, according to (9.21a),

$$\begin{aligned}
&\mathbb{P}_{\otimes X_i}(A_1 \times \cdots \times A_n) \\
&= \mathbb{P}\Big(\overset{n}{\underset{i=1}{\otimes}} X_i\Big)^*(A_1 \times \cdots \times A_n) \\
&= \mathbb{P}\Big[X_1^*(A_1) \cap \cdots \cap X_n^*(A_n)\Big]. \qquad (9.23a)
\end{aligned}$$

(ii) Suppose $X_1, \ldots, X_n$ are independent. Then, according to Problem 4.12,

$$\begin{aligned}
&\mathbb{P}_{\otimes X_i}(A_1 \times \cdots \times A_n) \\
&= \mathbb{P}\Big(\overset{n}{\underset{i=1}{\otimes}} X_i\Big)^*(A_1 \times \cdots \times A_n) \\
&= \mathbb{P}\Big[X_1^*(A_1) \cap \cdots \cap X_n^*(A_n)\Big] \\
&= \mathbb{P}_{X_1}(A_1) \cdots \mathbb{P}_{X_n}(A_n).
\end{aligned}$$

On the other hand, the product measure $\overset{n}{\underset{i=1}{\otimes}} \mathbb{P}_{X_i}$ has the factorization property on rectangles (9.23), meaning that $\overset{n}{\underset{i=1}{\otimes}} \mathbb{P}_{X_i}$ and $\mathbb{P}_{\otimes X_i}$ coincide on $\overset{n}{\underset{i=1}{\times}} \Sigma_i$. By Theorem 9.12 the two measures must be equal also on the product σ-algebra $\overset{n}{\underset{i=1}{\otimes}} \Sigma_i$. □

PROBLEMS

9.1 Let $\mathcal{G}_0$ be the set of all measurable rectangles $A = A_1 \times \cdots \times A_n$, $A_i \in \Sigma_i$. Denote by $\Sigma(\mathcal{G}_0)$ the algebra generated by $\mathcal{G}_0$ and by $\mathcal{C}(\mathcal{G}_0)$ the collection of all finite unions of disjoint rectangles of $\mathcal{G}_0$. Show that $\Sigma(\mathcal{G}_0) = \mathcal{C}(\mathcal{G}_0)$.

9.2 Prove that

$$\xi_{a_1}(A_1 \times A_2) = (A_1 \times A_2)_{a_1} = \begin{cases} A_2, & a_1 \in A_1 \\ \emptyset, & a_1 \notin A_1 \end{cases}.$$

9.3 Prove that the section is commutative with respect to all set operations.

9.4 Show the validity of assertion $a)$ in Lemma 9.9: *Let* σ_{a_i} *be* a_i*-sections for some* $a_i \in \Omega_i$, $i = 1, 2$. *Denote* $\Sigma' = \{A \in \Sigma_\otimes$: $\xi_{a_1}(A) = A_{a_1} \in \Sigma_2\}$. *Then* Σ' *is a* σ*-algebra in* $\Omega_1 \otimes \Omega_2$.

9.5 Let $A \subseteq \Omega_1 \times \Omega_2$ and let $a_1 \in \Omega_1$. Show that $(\mathbf{1}_A)_{a_1} = \mathbf{1}_{A_{a_1}}$, also in notation, $\varphi_{a_1} \circ \mathbf{1}_A = \mathbf{1}_{\varphi_{a_1}(A)}$.

9.6 Show that the Borel - Lebesgue measure λ^n on $\mathcal{B}^n$ and the product measure $\overset{n}{\underset{i=1}{\otimes}} \lambda_i$ on $\mathcal{B}_\otimes$ must be identical.

9.7 Let $(\Omega_i, \Sigma_i), i = 1, \ldots, n,$ be measurable spaces. Prove that

$$\overset{n}{\underset{i=1}{\otimes}} \Sigma_i = \left(\overset{n-1}{\underset{i=1}{\otimes}} \Sigma_i \right) \otimes \Sigma_n.$$

9.8 Let $(\Omega_i, \Sigma_i), i = 1, \ldots, n+k,$ be measurable spaces. Prove that

$$\overset{n+k}{\underset{i=1}{\otimes}} \Sigma_i = \left(\overset{n}{\underset{i=1}{\otimes}} \Sigma_i \right) \otimes \left(\overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i \right).$$

9.9 Let $(\mathbb{R}^n, \mathcal{B}^n, \lambda^n)$ and $(\mathbb{R}^k, \mathcal{B}^k, \lambda^k)$ be the Borel - Lebesgue measure spaces. Show that

$$(\mathbb{R}^n \times \mathbb{R}^k,\ \mathcal{B}^n \otimes \mathcal{B}^k,\ \lambda^n \otimes \lambda^k) = (\mathbb{R}^{n+k},\ \mathcal{B}^{n+k},\ \lambda^{n+k}).$$

9.10 Let $(\Omega_i, \Sigma_i, \mu_i),\ i = 1, 2,$ be measure spaces with σ-finite measures and let $A \in \Sigma_1 \otimes \Sigma_2$. Show that the following statements are equivalent.

1) $\mu_1 \otimes \mu_2(A) = 0$.

2) $\mu_2(A_{a_1}) = 0$ μ_1-a.e. on Ω_1.

3) $\mu_1(A_{a_2}) = 0$ μ_2-a.e. on Ω_2.

9.8 Let $A \subseteq \Omega_1 \times \Omega_2$ and let $a_1 \in \Omega_1$. Show that $(\mathbf{1}_A)_{a_1} = \mathbf{1}_{A_{a_1}}$.

9.9 Show that $f^*_{a_1}(A_3) = (f^*(A_3))_{a_1}$, $A_3 \subseteq \Omega_3$.

9.11 Let $(\Omega_i, \Sigma_i, \mu_i), i = 1, \ldots, n+k,$ be measurable spaces, such that μ_i's are σ-finite measures. Suppose $\bigotimes_{i=1}^{n} \mu_i$, $\bigotimes_{i=n+1}^{n+k} \mu_i$, and $\bigotimes_{i=1}^{n+k} \mu_i$ are product measures on product spaces $\bigotimes_{i=1}^{n} \Sigma_i$, $\bigotimes_{i=n+1}^{n+k} \Sigma_i$, and $\bigotimes_{i=n+1}^{n+k} \Sigma_i$, respectively. Prove that

$$\left(\bigotimes_{i=1}^{n} \mu_i\right) \otimes \left(\bigotimes_{i=n+1}^{n+k} \mu_i\right) = \bigotimes_{i=1}^{n+k} \mu_i.$$

9.12 Show that $f^*_{a_1}(A_3) = (f^*(A_3))_{a_1}$, $A_3 \subseteq \Omega_3$.

9.13 Prove Proposition 9.16. *Let $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$ be a measurable product space, (Ω_3, Σ_3) be a measurable space, and let $f : \Omega_1 \times \Omega_2 \to \Omega_3$ be a $\Sigma_1 \otimes \Sigma_2$-$\Sigma_3$-measurable function. Then the section f_{a_1} is Σ_2-Σ_3-measurable and f_{a_2} is Σ_1-Σ_3-measurable.* (*Hint*: Apply Lemma 9.9 and Problem 9.9.)

9.14 Show that $\xi_{a_1} \circ (\alpha f + \beta g) = \alpha \xi_{a_1} \circ f + \beta \xi_{a_1} \circ g$.

9.15 Let $f \in \mathcal{C}^{-1}_+(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$ and let a sequence of simple functions $\{s_n\} \subseteq \Psi_+(\Omega_1 \times \Omega_2, \mathcal{C}^{-1})$ be such that $f = \sup\{s_n\}$. Show that $f_{a_1} = \sup\{(s_n)_{a_1}\}$ (*Hint*: Apply Theorem 6.4, Chapter 5, and Problem 9.14.)

9.16 Show that $|f|_a = |f_a|$, $(f^+)_a = (f_a)^+$, and $(f^-)_a = (f_a)^-$.

9.17 Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be semi-rings on Ω_1 and Ω_2, respectively. Is $\mathcal{G}_1 \times \mathcal{G}_2$ also a semi-ring?

9.18 Let μ_i and ν_i be finite positive measures on measurable spaces (Ω_i, Σ_i), $i = 1, 2$. Show that if $\nu_i \ll \mu_i,\ i = 1, 2,$ then $\nu_1 \otimes \nu_2 \ll \mu_1 \otimes \mu_2$, where both product measures are on the product space $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$.

9.19 Let (Ω, Σ, μ) be a σ-finite measure space and let $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$. Prove that

$$\int f d\mu = \int_0^\infty \mu(\{f \geq x\})\lambda(dx).$$

(*Hint*: Use Theorem 9.11.)

9.20 Under the condition of Problem 9.19, let $g : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous monotone nondecreasing function such that $g(0) = 0$ and which is continuously differentiable on $(0, \infty)$. Show that

$$\int g(f) d\mu = (\mathrm{L}) \int_{(0,\infty)} g'(x)\mu(\{f \geq x\})\lambda(dx)$$

$$= (\mathrm{R}) \int_0^\infty g'(x)\mu(\{f \geq x\})dx.$$

9.21 Show that if F and G in Example 9.21 (vi) have no common discontinuities, then formula (9.21m) reduces to

$$F(b)G(b) - F(a-)G(a-) = \int_I F(x)\mu_G(dx) + \int_I G(x)\mu_F(dx). \tag{P9.21}$$

9.22 Show that the product map $\bigotimes_{i=1}^n X_i$ introduced in Remark 9.21 is Σ-$\bigotimes_{i=1}^n \Sigma_i$-measurable.

NEW TERMS:

Cauchy, Augustin-Louis, 498
Lebesgue, Henri, 498
Levi, Beppo, 498
Fubini, Guido, 498
Tonelli, Leonida, 498
Hobson, Ernest, 498
Valée-Poussin de la, Charles, 498
Euler, Leonard, 499
Lebesgue, Henri, 499
Radon, Johann, 499
$\overset{n}{\underset{i=1}{\times}} \Sigma_i$, measurable rectangles, 499
$\Sigma_{\otimes} = \overset{n}{\underset{i=1}{\otimes}} \Sigma_i = \Sigma\left(\overset{n}{\underset{i=1}{\times}} \Sigma_i\right)$, product σ-algebra, 499
section of a set, 504
$[\Sigma_{\otimes}, \mathcal{P}(\Omega_2), \sigma_{a_i}]$, a_i-section, 504
product measure, 510
product measure, uniqueness of, 510
$\varphi_{a_i} \circ f$, a_i-section of a function f, 513
Tonelli's theorem, 514
Fubini's theorem, 515
product measure space, 517
closed ball in $\mathbb{R}^n$, 517
Borel - Lebesgue measure of a closed ball in $\mathbb{R}^n$, 517
moment generating function of a random variable, 523
integration by parts formula for Lebesgue - Stieltjes integrals, 524, 531
$\left[\Omega, \prod_{i=1}^{n} \Omega_i, \overset{n}{\underset{i=1}{\otimes}} X_i\right]$, product map of random variables, 525-526
$\mathbb{P}_{\otimes X_i} = \mathbb{P}_{\overset{n}{\underset{i=1}{\otimes}} X_i} = \mathbb{P}\left(\overset{n}{\underset{i=1}{\otimes}} X_i\right)^*$, joint probability distribution, 527
$\left(\prod_{i=1}^{n} \Omega_i, \overset{n}{\underset{i=1}{\otimes}} \Sigma_i, \overset{n}{\underset{i=1}{\otimes}} \mathbb{P}_{X_i}\right)$, probability product space, 527
criterion of independence of random variables, 527

Chapter 7

Solutions to Selected Problems

Chapter 1. Set-Theoretic and Algebraic Preliminaries

1.5 Prove that:

$a)$ $(A\backslash B)^c = A^c \cup B$.
$b)$ $[(A^c \cup B)^c \cup (A \cup B^c)]^c = B\backslash A$.
$c)$ $(A \cap B) \cup (A \cap B^c) \cup (A^c \cap B) = A \cup B$.

Solution.

$a)$ Follows from $A\backslash B = A \cap B^c$ and the application of the idempotent and De Morgan's Laws.

$b)$ Using idempotent and De Morgan's laws we have

$$[(A^c \cup B)^c \cup (A \cup B^c)]^c = (A^c \cup B) \cap (A^c \cap B)$$

(by associative and distributive laws)

$$= [A^c \cup (B \cap A^c)] \cap B$$

(because $B \cap A^c$ is a subset of A^c)

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0_7, © Springer Science+Business Media, LLC 2013

$= B \cap A^c = B \backslash A$ (by Problem 1.3).

$c)$ $(A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$

$= (A \cap B) + (A \backslash B) + (B \backslash A) = A \cup B.$ □

1.6 For each of the following, justify with a proof or give a counterexample.

$a)$ $A \cup C = B \cup C \Rightarrow A = B.$
$b)$ $(A \cup B) \backslash B = A.$
$c)$ $A \backslash B = C \backslash B \Rightarrow A = C.$
$d)$ $A^c \backslash B^c = B \backslash A.$

Solution.

$a)$ Is not true. For instance, let A and B be arbitrary, but different subsets of C.

$b)$ Is not true: $(A \cup B) \backslash B = A \backslash B$.

$c)$ Is not true either. For example, let $A = B + C$ and none of the sets is empty. Then, $A \backslash B = C$ and $C \backslash B = C$. However, $A \neq C$. Thus, $A \backslash B = C \backslash B$.

$d)$ Is true : $A^c \backslash B^c = A^c \cap B = B \backslash A.$ □

1.9 Let $\{A_n : n = 1, 2, \dots\}$ be an arbitrary sequence of sets. Define $A_0 = \bigcap_{n=1}^{\infty} A_n$ and $A_\infty = \bigcup_{n=1}^{\infty} A_n$. Prove that

$$A_0 \subseteq \underline{\lim} A_n \subseteq \overline{\lim} A_n \subseteq A_\infty.$$

Solution. 1) Let $x \in \underline{\lim} A_n$. Then, by the definition, $x \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$. In other words, there is an n_0 such that $x \in \bigcap_{k=n_0}^{\infty} A_k$. Since $A_{n_0} \subseteq \bigcup_{k=1}^{\infty} A_k$, $A_{n_0+1} \subseteq \bigcup_{k=2}^{\infty} A_k, \ldots,$ it follows that also $x \in \bigcup_{k=1}^{\infty} A_k$, $x \in \bigcup_{k=2}^{\infty} A_k, \ldots,$ and hence, $x \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \overline{\lim} A_n$. Consequently, the middle inclusion holds.

2) Clearly, $A_0 \subseteq A_0 \cup \left(\bigcap_{k=2}^{\infty} A_k \right) \cup \left(\bigcap_{k=3}^{\infty} A_k \right) \cup \cdots = \underline{\lim} A_n$.

3) Obviously, $A_\infty \supseteq A_\infty \cap \left(\bigcup_{k=2}^{\infty} A_k \right) \cap \left(\bigcup_{k=3}^{\infty} A_k \right) \cap \cdots = \overline{\lim} A_n$.

□

1.10 Let Ω be an arbitrary set. Find a sequence $\{E_n\}$ of subsets of Ω such that

$$\underline{\lim}\, E_n = \emptyset \quad \text{and} \quad \overline{\lim}\, E_n = \Omega\,.$$

Solution. Let $E_{2n} = \Omega$ and $E_{2n-1} = \emptyset$. □

2.4 Show the equivalence of the following statements.

$a)$ f is injective.

$b)$ $f_*(A \cap B) = f_*(A) \cap f_*(B)$.

$c)$ For every pair A and B of disjoint sets, $f_*(A) \cap f_*(B) = \emptyset$.

What can go wrong with $b)$ if f is not injective?

Solution. We begin with some preliminaries. The problem is essentially based on Definition 2.1 (vii) of $f_*(A)$ and it can be interpreted as follows:

$$f_*(A) = \{y \in Y : \exists \text{ at least one } x \in A : f(x) = y\}. \qquad \text{(P2.4)}$$

Thus, for each $x \in A$, $f(x) \in f_*(A)$.

(i) We first show that $f_*(A \cap B) \subseteq f_*(A) \cap f_*(B)$ regardless of whether or not f is injective. Indeed, let $y \in f_*(A \cap B)$. Then, by (P2.4), there is at least one point $x \in A \cap B$ which is associated with y through f, i.e. $y = f(x)$. Then, $x \in A$ & $x \in B$ implies that $y = f(x) \in f_*(A) \cap f_*(B)$.

(ii) Now, assume that f is injective and show the inverse inclusion

$$f_*(A) \cap f_*(B) \subseteq f_*(A \cap B).$$

Let $y \in f_*(A) \cap f_*(B)$. Then, $y \in f_*(A)$ and $y \in f_*(B)$. By (P2.4), there is at least one $a \in A$ such that $y = f(a)$ and least one $b \in B$ such that $y = f(b)$. In general, a and b can be different, but because f injective, $a = b$, say x. Thus, $x \in A$ & $x \in B$ $\Rightarrow x \in A \cap B$ $\Rightarrow y = f(x) \in f_*(A \cap B)$.

Parts (i) and (ii) prove: $a) \Rightarrow b)$.

(iii) $b) \Rightarrow c)$ is trivial.

(iii) $c) \Rightarrow a)$. We show that $f(x_1) = f(x_2)$ implies $x_1 = x_2$. Let $A = \{x_1\}$ and $B = \{x_2\}$. Then, $A \cap B = \emptyset$, and by assumption $c)$, $f_*(\{x_1\}) \cap f_*(\{x_2\}) = \emptyset$.

(iv) Suppose f if not injective and let $y \in f_*(A) \cap f_*(B)$. Since $y \in f_*(A)$, there is an $a \in A$ such that $y = f(a) \in A$ and there is a $b \in B$ such that $y = f(b) \in B$, and because f is not injective, a need not equal b. If so, $\{a, b\}$ need not be a subset of $A \cap B$ and consequently,

$$\{y\} = f_*(\{a, b\}) \not\subseteq f_*(A \cap B).$$

The latter means that $y \notin f_*(A \cap B)$ and thus

$$f_*(A) \cap f_*(B) \not\subseteq f_*(A \cap B).$$

An easy counterexample illustrates the discrepancy of b). Let $f(x) = x^2$ and $A = (-3, 2)$ and $B = (1, 4)$. Then,

$$A \cap B = (1, 2),\ f_*(A) = [0, 9),\ f_*(B) = (1, 16),$$

$$f_*(A \cap B) = (1, 4), f_*(A) \cap f_*(B) = (1, 9). \qquad \square$$

In the following problems we assume that f is a map from $X = D_f$ into Y.

2.6 Show that $\forall B \subseteq Y,\ \ f_*(f^*(B)) \subseteq B$.

Solution. If $y \in f_*(f^*(B))$, then there is an $x \in X$ such that $x \in f^*(B)$ and $f(x) = y$. But $x \in f^*(B)$ implies that $y = f(x) \in B \cap Y \subseteq B$. $\square$

2.7 Show that $[X, Y, f]$ is surjective if and only if

$$f_*(f^*(B)) = B,\ \ \forall B \subseteq Y.$$

Solution.

(i) Let f be surjective and $B \subseteq Y$. Then, for $y \in B$, there is an $x \in X$ such that $f(x) = y$. (Notice that unless f is surjective, such an x need not exist, i.e. $f^*(\{y\})$ can be Ø.) Hence,

$$x \in f^*(\{y\}) \subseteq f^*(B) \ \Rightarrow\ x \in f^*(B)$$

$$\Rightarrow\ y = f(x) \in f_*(f^*(B)).$$

In as much as it holds for each $y \in B,\ B \subseteq f_*(f^*(B))$. The inverse inclusion follows from Problem 2.6 for all functions.

(ii) Let $f_*(f^*(B)) = B$. Then, in particular, for $B = \{y\}$,

$$f_*(f^*(\{y\})) = \{y\} \;\Rightarrow\; \forall y \in Y,\; f^*(\{y\}) \neq \emptyset$$

and thus f is onto. □

2.8 Show that the map $[\mathcal{P}(R_f), \mathcal{P}(D_f), f^*]$ is injective.

Solution. Suppose $B_1 \neq B_2 \in \mathcal{P}(R_f)$ and suppose without loss of generality that there is a $y_2 \in B_2$ such that $y_2 \notin B_1$. Then, obviously, sets $f^*(B_1)$ and $f^*(\{y_2\})$ are disjoint [or else there is an x belonging to both sets such that $f(x) \in B_1$ and $f(x) = y_2$ which means f is a multivalued function]. However, this is impossible if $f^*(B_1) = f^*(B_2)$. □

2.10 Is the map $[\mathcal{P}(R_f), \mathcal{P}(D_f), f^*]$ surjective?

Solution. No, it is not. Take for instance the map

$$[X = \{x_1, x_2\}, Y = \{y\}, f]$$

such that $f(x_1) = f(x_2) = y$. Then, $f^{**}(\mathcal{P}(Y)) = \{X, \emptyset\}$ which is a proper subset of $\mathcal{P}(X)$ and does not recover $\{x_1\}$ and $\{x_2\}$. □

3.1 Prove Theorem 3.1 under the condition that f is an into map. *Let* $[X, Y, f]$ *be an into map and let* $B \subseteq Y$. *Then*

$$[f^*(B)]^c = f^*(B^c).$$

Solution. Suppose $B \cap Y' \neq \emptyset$ and let B' denote the *trace* of B on Y'. By Theorem 3.1, applied to $[X, Y', f]$, we have that

$$[f^*(B')]^c = f^*((B')^c_{Y'}),$$

where B^c_A denotes $A \backslash B$. Obviously, $f^*(B') = f^*(B)$. Thus,

$$[f^*(B)]^c = [f^*(B')]^c = f^*((B')^c_{Y'})$$

and because $f^*(\{y\}) = \emptyset$ for all $y \in Y \backslash Y'$, the latter is also equal to $f^*(B^c)$. The case with $B \subseteq Y \backslash Y'$ is trivial. Indeed, $f^*(B) = \emptyset$ and $[f^*(B)]_X^c = X$. On the other hand, $Y' \subseteq B^c$, and

$$f^*(B^c) = f^*(Y') = X.$$

Combining the above arguments we have,

$$[f^*(B^c)]^c = X^c = \emptyset = f^*(B). \qquad \square$$

3.4 Let $[X, Y, f]$ be an into map and let $\{B_i : i \in I\}$ be an indexed family of subsets of Y.

a) Prove that $f^*(\bigcap_{i \in I} B_i) = \bigcap_{i \in I} f^*(B_i)$.

b) If B_1 and B_2 are disjoint, show that their inverse images under f are also disjoint.

c) If $\{B_i : i \in I\}$ is a pairwise disjoint family, show that

$$f^*(\textstyle\sum_{i \in I} B_i) = \sum_{i \in I} f^*(B_i).$$

Solution.

a) By the idempotent law,

$$f^*\left(\bigcap_{i \in I} B_i\right) = \left[\left\{f^*\left(\bigcap_{i \in I} B_i\right)\right\}^c\right]^c$$

by Problem 3.1 and De Morgan's law

$$= \left[f^*\left(\bigcup_{i \in I} B_i^c\right)\right]^c$$

by Problem 3.3

$$= \left(\bigcup_{i \in I} f^*(B_i^c)\right)^c$$

by Problem 3.1

$$= \left(\bigcup_{i \in I} [f^*(B_i)]^c \right)^c$$

and by De Morgan's and idempotent laws

$$= \bigcap_{i \in I} f^*(B_i).$$

b) Using a), $f^*(B_1) \cap f^*(B_2) = f^*(B_1 \cap B_2) = f^*(\emptyset) = \emptyset$. The latter is due to Example 3.2.

c) $f^*(\sum_{i \in I} B_i) = \bigcup_{i \in I} f^*(B_i)$ holds true for an arbitrary sequence of sets under the union, due to Problem 3.3. That the right-hand side equals $\sum_{i \in I} f^*(B_i)$ is due to b), because $\{f^*(B_i)\}$ is a family of pairwise disjoint sets. □

3.6 The results above prove that all set operations are closed under the inverses of maps. Show that not all set operations are closed under maps as per the following.

a) Show that maps preserve inclusions.

b) Show that maps preserve unions.

c) Show that maps do not preserve intersections; specifically, show that

$$f_* \left(\bigcap_{i \in I} A_i \right) \subseteq \bigcap_{i \in I} f_*(A_i)$$

and that the inverse inclusion need not hold. Explain the latter with and without a counterexample.

d) Do maps preserve the difference?

Solution.

a) Let $A \subseteq B$ and let $y \in f_*(A)$. Then, there is an $x \in A$ such that $f(x) = y$. From $A \subseteq B$, it follows that $x \in B$ and thus $y = f(x) \in f_*(B)$. Because $y = f(x)$ was assumed to belong to $f_*(A)$ and $y \in f_*(B)$, we have the pledged inclusion.

Here is another proof.

Let $A \subseteq X$ ($= D_f$ without loss of generality).

(i) Let $y \in f_*(A)$. Then, obviously $f^*(\{y\}) \cap A \neq \emptyset$.

(ii) If $y \notin f_*(A)$, then, $f^*(\{y\}) \cap A = \emptyset$.

Now, if given $A \subseteq B$, $f_*(A) \nsubseteq f_*(B)$, then there is at least one point $y \in f_*(A)$ that does not belong to $f_*(B)$. Now, this leads to a contradiction due to (i) and (ii), in as much as $f^*(\{y\})$ intersects A but not B.

[Notice that if f is inconsistent, then the worst case scenario is that $f_*(A)$ maybe empty, but then the inclusion $f_*(A) \subseteq f_*(B)$ automatically holds.]

b) We show that for any indexed family $\{A_i;\, i \in I\}$,

$$f_*\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f_*(A_i).$$

(i) If $y \in f_*\left(\bigcup_{i\in I} A_i\right)$, there is an $x \in \bigcup_{i\in I} A_i$ such that $f(x) = y$. Furthermore, there is an $i_0 \in I$ such that $x \in A_{i_0}$ and hence,

$$y = f(x) \in f_*(A_{i_0}) \subseteq \bigcup_{i\in I} f_*(A_i).$$

Consequently,

$$f_*\left(\bigcup_{i\in I} A_i\right) \subseteq \bigcup_{i\in I} f_*(A_i).$$

(ii) If $y \in \bigcup_{i\in I} f_*(A_i)$, then there is an $i_0 \in I$ such that $y \in f_*(A_{i_0})$ and thus there is an x such that $f(x) = y \in f_*\left(\bigcup_{i\in I} A_i\right)$. This yields

$$f_*\left(\bigcup_{i\in I} A_i\right) \supseteq \bigcup_{i\in I} f_*(A_i).$$

$c)$ (i) If $y \in f_*\left(\bigcap_{i\in I} A_i\right)$, then there is an $x \in \bigcap_{i\in I} A_i$ such that $f(x) = y$. Thus, x belongs to each A_i and $y = f(x)$ belongs to each $f_*(A_i)$. The latter yields that $y \in \bigcap_{i\in I} f_*(A_i)$ and hence

$$f_*\left(\bigcap_{i\in I} A_i\right) \subseteq \bigcap_{i\in I} f_*(A_i).$$

(ii) Let $y \in \bigcap_{i\in I} f_*(A_i)$. Then for each i, $y \in f_*(A_i)$ and hence, for each $i \in I$, there is at least one point, say $x_i \in A_i$, such that $f(x_i) = y$. Let $A_i' = f^*(\{y\}) \cap A_i$. Now, if $\bigcap_{i\in I} A_i' \neq \emptyset$, then for each $x \in \bigcap_{i\in I} A_i'$,

$$f(x) = y \in f_*\left(\bigcap_{i\in I} A_i'\right) \subseteq f_*\left(\bigcap_{i\in I} A_i\right)$$

implying that $\bigcap_{i\in I} f_*(A_i) \subseteq f_*\left(\bigcap_{i\in I} A_i\right)$. However, if $\bigcap_{i\in I} A_i' = \emptyset$, there is no $x \in \bigcap_{i\in I} A_i$ such that $f(x) = y$ and thus $\bigcap_{i\in I} f_*(A_i) \nsubseteq f_*\left(\bigcap_{i\in I} A_i\right)$. Exactly it means that for at least one $y \in \bigcap_{i\in I} f_*(A_i)$, all of its inverse images under f are scattered over A_i's, with not a single point common for all.

As far as a counterexample, we can take $[\mathbb{R}, \mathbb{R}_+, f(x) = x^2]$. Let $A = [-2, -1]$ and $B = [1, 2]$. Obviously, these two sets are disjoint, but not their images under f which are identically $[1, 4]$.

d) Let $f = \mathbf{1}_{[0,4]}$, $A = [0,3]$ and $B = (1,5]$. Then, $A\backslash B = [0,1]$ and thus $\mathbf{1}_{[0,4]}([0,1]) = \{1\}$, while $\mathbf{1}_{[0,4]}(A)\backslash\mathbf{1}_{[0,4]}(B) = \{1\}\backslash\{0,1\} = \emptyset$ implying that the difference under this f is not preserved. □

3.9 Let $\{A_n\}$ be a sequence of subsets of Ω. Show that the function $\overline{\lim} 1_{A_n}$ is the indicator function of the set $\overline{\lim}\, A_n$ (i.e., $\mathbf{1}_{\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m}$) and that the function $\underline{\lim}\, 1_{A_n}$ is the indicator function of the set $\underline{\lim}\, A_n$ (i.e., $\mathbf{1}_{\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m}$).

Solution. (i) We need to show that $\overline{\lim}\mathbf{1}_{A_n} = \mathbf{1}_{\overline{\lim} A_n}$. Recall that

$$\overline{\lim} A_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k.$$

If $x \in \overline{\lim} A_n$, then $\mathbf{1}_{\overline{\lim} A_n}(x) = 1$ and

$$x \in \bigcup_{k=1}^{\infty} A_k, x \in \bigcup_{k=2}^{\infty} A_k, \ldots \tag{P3.9}$$

Therefore, there is an n_1 such that $x \in A_{n_1}$. From (P3.9) it follows that $x \in \bigcup_{k=n_1+1}^{\infty} A_k$. Hence, there is an $n_2 > n_1$ such that $x \in A_{n_2}$. Continuing with this process, we arrive at infinite subsequence $\{A_{n_j}\}$ such that x belongs to all of these sets. Consequently, $\mathbf{1}_{A_{n_j}}(x) = 1, j = 1, 2, \ldots$. As far as the rest of A_k, they are zeros or ones. Hence $\overline{\lim}\mathbf{1}_{A_n} = 1$.

Now, let $x \notin \overline{\lim}\, A_n$. Then, $\mathbf{1}_{\overline{\lim} A_n}(x) = 0$ and $x \notin \bigcup_{k=1}^{\infty} A_k, x \notin \bigcup_{k=2}^{\infty} A_k, \ldots$ implying that $x \notin A_n$ for all $n = 1, 2, \ldots$. So, $\mathbf{1}_{A_n}(x) = 0$ for all n and thus we have $\overline{\lim}\mathbf{1}_{A_n}(x) = 0$.

(ii) We show that $\underline{\lim}\, \mathbf{1}_{A_n} = \mathbf{1}_{\underline{\lim} A_n}$. Let

$$x \in \underline{\lim}\, A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Then $\mathbf{1}_{\underline{\lim} A_n}(x) = 1$. Furthermore, there is at least one index n_0 such that point $x \in \bigcap_{k=n_0}^{\infty} A_k$ which implies that

$$\mathbf{1}_{A_{n_0}}(x) = \mathbf{1}_{A_{n_0+1}}(x) = \cdots = 1.$$

Consequently, the whole tail of the sequence $\{\mathbf{1}_{A_n}(x)\}$ equals 1 and thus $\lim_{n\to\infty} \mathbf{1}_{A_n}(x) = 1$.

If $x \notin \underline{\lim}\, A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$, then $\mathbf{1}_{\underline{\lim} A_n}(x) = 0$. On the other hand, by De Morgan's law, $x \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k^c$ and hence, by part (i), $\mathbf{1}_{A_{n_j}}(x) = 0$, $j = 1, 2, \ldots$, that is, there is a subsequence of $\{\mathbf{1}_{A_n}(x)\}$ convergent to zero. Hence $\underline{\lim}\, \mathbf{1}_{A_n}(x) = 0$. □

3.12 Let $[X, Y, f]$ be a consistent map. Denote $F = f_* \circ f^*$. Then, the map $[\mathcal{P}(Y), \mathcal{P}(Y), F]$ is such that $F = I_*$ (I is the identity map) if and only if $[X, Y, f]$ is surjective.

Solution. (i) Let f be surjective. Then, for each $y \in Y$,

$$\emptyset \neq f^*(\{y\}) = A \subseteq X \quad \text{such that} \quad f_*(A) = \{y\}.$$

This shows that $F\{y\} = f_* \circ f^*(\{y\}) = I_*\{y\} = \{y\}$. [Otherwise, if $\{y\}$ is a proper subset of $f_*(A)$, then there is at least one $x \in A$ such that $f(x) = y$ and $f(x) = z \neq y$.]

(ii) Let $B \subseteq Y$. Then, $B = \bigcup_{y\in B} \{y\}$ and because the inverse preserves all set operations, we have

$$f^*(B) = \bigcup_{y\in B} f^*(\{y\}).$$

Now in as much as f_* preserves unions,

$$f_*(f^*(B)) = \bigcup_{y \in B} f_*(f^*(\{y\}))$$

which by (i) equals $\bigcup_{y \in B} \{y\} = B$. This completes the proof that $f_* \circ f^* = I_*$ if f is onto.

(iii) Suppose that f is not surjective. Then there is a subset B of Y such that $B = B_1 + B_2$ and $B_1 \subseteq R_f$, and $B_2 \subseteq Y \backslash R_f$. Now,

$$f_* \circ f^*(B) = f_*(f^*(B_1) + \emptyset) = B_1 \neq B$$

by (ii) because $B_1 \subseteq R_f$. □

4.2 Classify the following binary relations.
$a)$ Let Ω be a nonempty set. Define the relation $(\mathcal{P}(\Omega), \subseteq)$.
$b)$ Let $\Omega = \mathbb{R}^2 \backslash (x, 0)$. Define R: $(a, b)\mathrm{R}(c, d) \Leftrightarrow ad = bc$.

Solution.

$a)$ Partial order.

$b)$ The collection of all points of Ω involved in each relation are those on each individual line through the origin (excluding the origin itself). This argument is based on similarity of right triangles with vertices on these lines, on the line-by-line basis. Thus, all points (excluding the origin) in each individual line are related and only they. Thereby, R partitions Ω into disjoint lines through the origin and R is an equivalence relation. □

4.3 The following theorem is a statement of the *principle of mathematical induction.*

Let $S(n)$ be a statement which is true or false, for $n = 1, 2, \ldots$. Let $S(1)$ be true and suppose whenever $S(n)$ is true

implies that $S(n+1)$ *is true,* $n = 1, 2, \ldots$ *. Then* $S(n)$ *is true for all* n*.*

Prove it. (*Hint*: Use the well-ordering principle.)

4.6 Let l^0 denote the space of all summable real-valued sequences. In other words, a sequence $[\mathbb{N}, \mathbb{R}, f]$, in the form of $\{a_n\}$, is *summable*, if the series $\sum_{n=1}^{\infty} a_n$ converges. We define the partial order $\preceq$ on l^0 as $f \preceq g$ if and only if $f(n) \leq g(n)$ for all $n \in \mathbb{N}$. Is $(l^0, \preceq)$ a lattice?

Solution. No, it is not. Take $f, g \in l^0$ such that $f(n) = (-1)^n \frac{1}{n}$ and $g(n) = (-1)^{n+1} \frac{1}{n}$. Then, either of these sequences is summable, because their associated series are Leibnitz. However,

$$f \vee g = \left\{1, \tfrac{1}{2}, \tfrac{1}{3}, \ldots\right\}$$

is clearly not summable. □

4.8 Is the set of all continuous real-valued functions a lattice?

Solution. From $(f \vee g) = \frac{1}{2}(f + g + |f - g|)$ and $(f \wedge g) = \frac{1}{2}(f + g - |f - g|)$, we see that $(f \vee g)$ and $(f \wedge g)$ are continuous. Thus, the set of all continuous functions is a lattice. □

4.10 Prove Theorem 4.4: *Let* $[X, Y, f]$ *be a consistent function with* E_f *being its equivalence kernel. Then there is a unique function* $[X|\mathrm{E}_f, Y, \boldsymbol{f}]$ *such that* $f = \boldsymbol{f} \circ \pi_{\mathrm{E}_f}$.

Solution. Let $x_1, x_2 \in X$. An element $x_1 \in [x_2]_{\mathrm{E}_f}$ if and only if $x_1 \mathrm{E}_f x_2$ [i.e., if and only if $\pi_{\mathrm{E}_f}(x_1) = \pi_{\mathrm{E}_f}(x_2)$]. Therefore, $\pi_{\mathrm{E}_f}(x_1)$ is the only equivalence class that contains x_1. Consequently, we showed that each element of X belongs to exactly one equivalence class for E_f, even if E_f is merely an arbitrary equivalence relation on X. Now, because $x_1 \mathrm{E}_f x_2$ implies $f(x_1) = f(x_2)$ [see Example 4.2 $(viii)$], the function f assigns all elements x of an equivalence class $\pi_{\mathrm{E}_f}(x)$ to a single element y of Y. Therefore, $\boldsymbol{f}$ can be defined as a

map assigning the entire class $[x]_{\mathrm{E}_f} = \pi_{\mathrm{E}_f}(x)$ to a single element $y = f(x)$ for all $x \in \pi_{\mathrm{E}_f}(x)$. The correspondence $\pi_{\mathrm{E}_f}(x) \mapsto y$ is therefore a function $\boldsymbol{f}$: $X|\mathrm{E}_f \to Y$ such that $f = \boldsymbol{f} \circ \pi_{\mathrm{E}_f}$. Obviously, such a function is unique. □

4.12 Prove part (b) of Corollary 4.5: *Let* $[X, Y, f]$ *be a function and let* E_f *denote its equivalence kernel. Then if* f *is surjective*, *the reducer* $\boldsymbol{f}$ *is bijective*.

Solution. By the assumption, $f_*(X) = Y$ implying that for each $y \in Y$, there is an $x \in X$ such that $f(x) = y$. Thus, for each $y \in Y$, there is a nonempty class $[x]_{\mathrm{E}_f}$ and consequently, $\boldsymbol{f}_*(X|E_f) = Y$, that is, $\boldsymbol{f}$ is surjective. From part (a) of Corollary 4.5, $[X|\mathrm{E}_f, Y, \boldsymbol{f}]$ injective, which proves the assertion. □

5.1 For subsets $A \subseteq Y_1$ and $B, C \subseteq Y_2$, prove that

(i) $\mathbf{1}_{A\times B}(x, y) = \mathbf{1}_A(x)\mathbf{1}_B(y), \forall x \in A, y \in B.$

(ii) $(A \times B)^c = (A^c \times Y_2) \cup (Y_1 \times B^c).$

Show the validity of the distributive law with respect to cross product and union, intersection, and difference:

(iii) $A \times (B \cup C) = (A \times B) \cup (A \times C).$

(iv) $A \times (B \cap C) = (A \times B) \cap (A \times C).$

(v) $A \times (B \backslash C) = (A \times B) \backslash (A \times C).$

Solution.

(i) Show that the left-hand side and right-hand side are simultaneously equal to 1 or 0 dependent on whether $(x, y) \in A \times B$.

(ii) $(A \times B)^c = [\pi_1^*(A) \cap \pi_2^*(B)]^c = \pi_1^*(A^c) \cup \pi_2^*(B^c)$

$= (A^c \times Y_2) \cup (Y_1 \times B^c).$

For problems $(iii\text{-}v)$ we use the pick-a-point process or inverse projections:

$$\begin{aligned}(iii)\ \ (a,\omega)\in A\times(B\cup C) &\Leftrightarrow (a\in A)\wedge(\omega\in B\cup C)\\ &\Leftrightarrow (a\in A)\wedge(\{\omega\in B\}\vee\{\omega\in C\})\\ &\Leftrightarrow (\{a\in A\}\wedge\{\omega\in B\})\vee(\{a\in A\}\wedge\{\omega\in C\})\\ &\Leftrightarrow ((a,\omega)\in A\times B)\vee((a,\omega)\in A\times C)\\ &\Leftrightarrow (a,\omega)\in(A\times B)\cup(A\times C).\end{aligned}$$

Alternatively,

$$A\times(B\cup C)=\pi_1^*(A)\cap\pi_2^*(B\cup C)=\pi_1^*(A)\cap[\pi_2^*(B)\cup\pi_2^*(C)]$$

$$[\pi_1^*(A)\cap\pi_2^*(B)]\cup[\pi_1^*(A)\cap\pi_2^*(C)]=(A\times B)\cup(A\times C).$$

(iv) Similar to (ii) if $\vee$ is replaced with $\wedge$:

$$\begin{aligned}&(a,\omega)\in A\times(B\cap C)\Leftrightarrow(a\in A)\wedge(\omega\in B\cap C)\\ &\Leftrightarrow(a\in A)\wedge(\{\omega\in B\}\wedge\{\omega\in C\})\\ &\Leftrightarrow(\{a\in A\}\wedge\{\omega\in B\})\wedge(\{a\in A\}\wedge\{\omega\in C\})\\ &\Leftrightarrow((a,\omega)\in A\times B)\wedge((a,\omega)\in A\times C)\\ &\Leftrightarrow(a,\omega)\in(A\times B)\cap(A\times C).\end{aligned}$$

(v) Use the equation $B\backslash C=B\cap C^c$ and then apply (iii):

$$\begin{aligned}&(a,\omega)\in A\times(B\backslash C)\Leftrightarrow(a\in A)\wedge(\omega\in B\cap C^c)\\ &\Leftrightarrow(a\in A)\wedge(\{\omega\in B\}\wedge\{\omega\in C^c\})\\ &\Leftrightarrow(\{a\in A\}\wedge\{\omega\in B\})\wedge(\{a\in A\}\wedge\{\omega\in C^c\})\\ &\Leftrightarrow((a,\omega)\in A\times B)\wedge((a,\omega)\in A\times C^c)\\ &\Leftrightarrow((a,\omega)\in A\times B)\wedge((a,\omega)\notin A\times C)\\ &\Leftrightarrow(a,\omega)\in(A\times B)\backslash(A\times C).\end{aligned}$$

Here is why. Formally,

$$(a,\omega) \notin A \times C \Leftrightarrow (a \notin A \wedge \omega \in C) \vee (a \in A \wedge \omega \in C^c).$$

But the combination of $(a \notin A \wedge \omega \in C)$ with $(a,\omega) \in A \times B$ is impossible. □

5.3 Is it true that

$$(A_1 \cap A_2) \times (B_1 \cap B_2) = (A_1 \times B_2) \cap (A_2 \times B_1)?$$

Solution. Yes it is true. It can be easily shown that

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \times B_2) \cap (A_2 \times B_1).$$ □

5.4 On the other hand, using the distributive law in Problem 5.1 (iii) it appears that

$$\begin{aligned}&(A_1 \cap A_2) \times (B_1 \cap B_2)\\ &= (A_1 \times B_1) \cap (A_2 \times B_2)\\ &\cap (A_1 \times B_2) \cap (A_2 \times B_1).\end{aligned}$$

Is this correct?

Solution. Yes, it is correct because

$$(A_1 \cap A_2) \times (B_1 \cap B_2) = (A_1 \times B_2) \cap (A_2 \times B_1),$$

due to Problem 5.3. □

5.6 Prove that

(i) $\quad A \times \bigcup_{i\in I} B_i = \bigcup_{i\in I} (A \times B_i).$

(ii) $\quad \left[\bigcup_{i\in I} A_i\right] \times B = \bigcup_{i\in I} (A_i \times B).$

$(iii)\ \ A \times \bigcap_{i\in I} B_i = \bigcap_{i\in I} (A \times B_i).$

Solution. $(ii)\ \left[\bigcup_{i\in I} A_i\right] \times B = \pi_1^*\left(\bigcup_{i\in I} A_i\right) \cap \pi_2^*(B)$

$$= \left[\bigcup_{i\in I} \pi_1^*(A_i)\right] \cap \pi_2^*(B) = \bigcup_{i\in I} \{\pi_1^*(A_i) \cap \pi_2^*(B)\}$$

$$= \bigcup_{i\in I} (A_i \times B).$$

(iii) Follows by the distributive law and induction. □

5.9 Let $A = A_1 \times A_2$ and $B = B_1 \times B_2$ be subsets of $Y_1 \times Y_2$ such that $A_i, B_i \subseteq Y_i, i = 1, 2$. Prove the validity of the decomposition

$$A\backslash B = [(A_1\backslash B_1) \times A_2] + [(A_1 \cap B_1) \times (A_2\backslash B_2)].$$

In other words, $A\backslash B$ can be partitioned as two disjoint "rectangles: $(A_1\backslash B_1) \times A_2$ and $(A_1 \cap B_1) \times (A_2\backslash B_2)$.

Show the validity of yet another decomposition of $A\backslash B$:

$$A\backslash B = [A_1 \times (A_2\backslash B_2)] + [(A_1\backslash B_1) \times (A_2 \cap B_2)].$$

Solution. Notice that from Problem 5.2(ii) we have

$$A\backslash B = (A_1 \times A_2)\backslash(B_1 \times B_2)$$

$$= [(A_1\backslash B_1) \times A_2] \cup [A_1 \times (A_2\backslash B_2)]. \qquad \text{(P5.9)}$$

In (P5.9), the underlying sets are overlapping. However, we utilize (P5.9) to get the expression

$$A\backslash B = [(A_1\backslash B_1) \times A_2] + [(A_1 \cap B_1) \times (A_2\backslash B_2)]. \qquad \text{(P5.9a)}$$

Clearly, because of (P5.9),

$$[(A_1\backslash B_1)\times A_2]+[(A_1\cap B_1)\times(A_2\backslash B_2)]\subseteq A\backslash B.$$

Next, the sets $[(A_1\backslash B_1)\times A_2]$ and $[(A_1\cap B_1)\times(A_2\backslash B_2)]$ are disjoint, because so are $A_1\backslash B_1$ and $A_1\cap B_1$.

We need to prove the inverse inclusion. If $\omega\in[(A_1\backslash B_1)\times A_2]$, then this is because the latter is shared by both representations in (P5.9) and (P5.9a). So, it follows that ω also belongs to the right-hand side of (P5.9a). Therefore, assume that $\omega=(x,y)\in A_1\times(A_2\backslash B_2)$ [the second set in (P5.9)]. Then, $x\in A_1$ and $y\in A_2\backslash B_2$. If, in addition, x also belongs to B_1, then $x\in A_1\cap B_1$ and thus, ω belongs to the second term on the right of (P5.9a): $(A_1\cap B_1)\times(A_2\backslash B_2)$. Otherwise, if $x\notin B_1$, then $x\in A_1\backslash B_1$ and thus ω belongs to the first term on the right in (P5.9a). Either way, ω is an element of the right-hand side of (P5.9a) showing the validity of the inverse inclusion.

Therefore, from (P5.9a) it follows that $A\backslash B$ is the union of two disjoint rectangles.

The other decomposition of $A\backslash B$,

$$A\backslash B=[A_1\times(A_2\backslash B_2)]+[(A_1\backslash B_1)\times(A_2\cap B_2)]$$

can be shown similarly or by interchanging the roles of A_i's and B_i's. □

5.10 Let $\{A_x\subseteq Y_x;\ x\in X\}$ and $\{Z_y\subseteq Y_y;y\in X\}$ be two family of subsets. Show that

a) $$\left(\prod_{x\in X}A_x\right)\cap\left(\prod_{y\in X}B_y\right)=\prod_{x\in X}(A_x\cap B_x).$$

b) $$\left(\prod_{x\in X}A_x\right)\cup\left(\prod_{y\in X}B_x\right)\supseteq\prod_{x\in X}(A_x\cap B_x).$$

Solution.

$$a)\ \left(\prod_{x\in X} A_x\right)\cap\left(\prod_{y\in X} B_y\right)=\left[\bigcap_{x\in X}\pi_x^*(A_x)\right]\cap\left[\bigcap_{y\in X}\pi_y^*(B_y)\right]$$

(because the inverse of any map preserves the intersection)

$$=\bigcap_{x\in X}(\pi_x^*(A_x)\cap\pi_x^*(B_x))$$

$$=\bigcap_{x\in X}\pi_x^*(A_x\cap B_x)=\prod_{x\in X}(A_x\cap B_x).$$

$b)$ Follows from $a)$ and $A\cap B\subseteq A\cup B$. □

5.12 Let $\{A_i : i\in I\}$ and $\{B_j : j\in J\}$ be two indexed families of subsets of Y_1 and Y_2, respectively. Show that

$$\left[\bigcap_{i\in I} A_i\right]\times\left[\bigcap_{j\in J} B_j\right]=\bigcap_{i\in I}\bigcap_{j\in J}(A_i\times B_j).$$

Solution.

$$\left[\bigcap_{i\in I} A_i\right]\times\left[\bigcap_{j\in J} B_j\right]$$

$$=\pi_1^*\left[\bigcap_{i\in I} A_i\right]\cap\pi_2^*\left[\bigcap_{j\in J} B_j\right]$$

$$=\bigcap_{i\in I}\pi_1^*(A_i)\cap\left[\bigcap_{j\in J}\pi_2^*(B_j)\right]$$

$$=\bigcap_{i\in I}\bigcap_{j\in J}(\pi_1^*(A_i)\cap\pi_2^*(B_j))=\bigcap_{i\in I}\bigcap_{j\in J}(A_i\times B_j).$$ □

6.3 (a) Let A be an uncountable set and let $B\subseteq A$ be countable. Show that $A\backslash B$ is uncountable. (b) If $|A|=\mathfrak{c}$ and $|B|\leq\aleph_0$, show that $|A\backslash B|=\mathfrak{c}$.

Solution.

(a) If $A\backslash B$ were at most countable, then $A = B + (A\backslash B)$ is at most countable, which is a contradiction.

(b) We have

$$\mathbb{N} \preceq A\backslash B \preceq A.$$

By $a)$, $A\backslash B$ is uncountable and thus we have:

$$\aleph_0 < |A\backslash B| \leq \mathfrak{c}.$$

Therefore $|A\backslash B| = \mathfrak{c}$, which follows from the continuum hypothesis [Remark 6.5(ii)]. □

6.6 Show that $2^{\aleph_0} = \mathfrak{c}$.

Solution. Define $f\colon \mathbb{R} \to \mathcal{P}(\mathbb{Q})$ (denoting the latter for notational convenience by $2^{\mathbb{Q}}$) as $f(x) = \{q \in \mathbb{Q}\colon q < x\}$, $x \in \mathbb{R}$, and $\mathbb{Q}$ denotes the set of all rational numbers.

$a)$ We show that f is injective. Let $a < b \in \mathbb{R}$. Then there is a $q \in \mathbb{Q}$ such that $a < q < b$. By the above definition, $q \in f(b)$ and $q \notin f(a)$. Therefore, $f(a) \neq f(b)$. The same holds if $b < a$. Consequently, $\mathbb{R} \preceq \mathcal{P}(\mathbb{Q})$.

$b)$ Let $I(\mathbb{N})$ denote the set of all indicator functions on $\mathbb{N}$. As we know, $I(\mathbb{N}) \approx 2^{\mathbb{N}}$. Let $F\colon I(\mathbb{N}) \to [0, 1]$ such that

$$F(\mathbf{1}_A) = 0.\mathbf{1}_A(1)\mathbf{1}_A(2)\cdots$$

If $A \neq B$, then $\mathbf{1}_A \neq \mathbf{1}_B$ and thus $F(\mathbf{1}_A) \neq F(\mathbf{1}_B)$, which implies that F is injective and then

$$I(\mathbb{N}) \approx 2^{\mathbb{N}} \preceq [0, 1].$$

But $2^{\mathbb{N}} \approx 2^{\mathbb{Q}}$ and $[0, 1] \approx \mathbb{R}$ (see Problem 6.5). Hence $2^{\mathbb{N}} \preceq \mathbb{R}$, whereas from $a)$, $\mathbb{R} \preceq 2^{\mathbb{N}}$. By the Schröder - Bernstein theorem (Problem 6.2), $2^{\mathbb{N}} \approx \mathbb{R}$ or $2^{\aleph_0} = \mathfrak{C}$. □

6.7 Let $[X, Y, f]$ be a surjective map. Show that there is a subset of X equipotent with Y.

Solution. By the axiom of choice, for each $y \in Y$, there is a point $x_y \in f^*(\{y\})$. Let $X' = \bigcup_{y \in Y} \{x_y\}$. Then, $[X', Y, \mathrm{Res}_{X'} f]$ is bijective and thus $X' \approx Y$. □

6.9 We call an *algebraic number* any root of a polynomial with integer coefficients. What is the cardinal number of all algebraic numbers?

Solution. Let $\mathcal{P}_n$ be the set of all n-degree polynomials with integer coefficients. Then, clearly, $\mathcal{P}_n \approx \mathbb{Z}^{n+1}$ ($n+1$ times the Cartesian product of all integer numbers). Each such polynomial has at most n distinct zeros. If $\mathcal{S}(\mathcal{P}_n)$ is the set of all algebraic numbers, then it is $\preceq$ -related as follows.

$$\mathbb{N} \preceq \mathcal{S}(\mathcal{P}_n) \preceq N \times \mathbb{Z}^{n+1} \approx \mathbb{N},$$

where $N = \{0, 1, \ldots, n\}$. By the Schröder-Bernstein theorem, $\mathcal{S}(\mathcal{P}_n) \approx \mathbb{N}$. If $\mathcal{S}(\mathcal{P})$ is the set of all algebraic numbers, then

$$\mathcal{S}(\mathcal{P}) = \bigcup_{n=1}^{\infty} \mathcal{S}(\mathcal{P}_n)$$

and the latter has the same cardinal number as $\mathbb{N}$, according to Example 6.3(iv). □

6.10 Prove the statement. *Every infinite set contains a countable subset.*

Solution. Let A be an infinite set. Denote $f(1) = a_1 \in A$, $f(2) = a_2 \in A \backslash \{a_1\}, \ldots$. We can use Problem 6.7 each time to warrant the availability of elements in $A \backslash \{a_1, \ldots, a_n\}$. Therefore, $[\mathbb{N}, A, f]$ is injective. Consequently, $\mathbb{N} \preceq A$ and $\mathbb{N} \approx f(\mathbb{N}) \subseteq A$. □

6.14 Let $\mathfrak{A} = \{B_x : x \in X\}$ be a countable family (i.e., $X \approx \mathbb{N}$) of countable sets. Let

$$\mathcal{B} = \{B_{x_1} \times \cdots \times B_{x_n} : n \in \mathbb{N}, A = \{x_1, \ldots, x_n\} \subseteq X\}$$

denote the set of all finite Cartesian products from $\mathfrak{A}$. Show that $\mathcal{B}$ is countable.

Solution. If $n \in \mathbb{N}$ is chosen, then $A = \{x_1, \ldots, x_n\} \subseteq X$ is selected, along with $B_{x_1} \times \cdots \times B_{x_n}$, then finally a point $\mathbf{b}_n = \{b_{x_1}, \ldots, b_{x_n}\}$ (or rather an n-element set) from set $B_{x_1} \times \cdots \times B_{x_n}$ is picked out.

In general, all n-element sets can be picked out arbitrarily from B_i's, but according to the above Cartesian order, each one of them has at most one element from any B_i.

Now, let $\mathfrak{X}$ be the set of all finite subsets of X. Then, picking out a set from $\mathfrak{X}$, such as $\{x_1, \ldots, x_n\}$, we then form the Cartesian product such as $B_{x_1} \times \cdots \times B_{x_n}$ and then select all elements $\mathbf{b}_n$ from the set $B_{x_1} \times \cdots \times B_{x_n}$ to include them in $\mathcal{B}$.

Now, if we organize elements of $\mathfrak{A}$ in an infinite matrix (assuming for a moment that $X = \mathbb{N}$), with B_x's being its columns, then by Example 6.3(iv), this matrix is enumerable, that is, the matrix is countable. The set of all of its finite subsets, according to Problem 6.13, must be countable. Let $\mathfrak{A}_n$ be the set of all subsets of this matrix having n elements and let $\mathfrak{B}_n$ be the set of all n element sets selected in accordance with the above specification, such as set $\mathbf{b}_n$. Then clearly, $\mathfrak{B}_n \subseteq \mathfrak{A}_n$ because $\mathfrak{B}_n$ by its nature is restrictive, whereas $\mathfrak{A}_n$ is not. For example, $\mathfrak{A}_n$ may contain some subsets with all elements chosen from one and the same set B_x. Consequently,

$$\mathcal{B} = \bigcup_{n=1}^{\infty} \mathfrak{B}_n \subseteq \bigcup_{n=1}^{\infty} \mathfrak{A}_n \approx \mathbb{N}. \qquad \square$$

7.1 Show that each monoid has exactly one identity.

Solution. Let $(\mathcal{G}, *)$ be a monoid with two identities, I_1 and I_2. Then, $I_1 * I_2 = I_1$ and $I_2 * I_1 = I_2$. Because $*$ is commutative with respect to an identity, the above two equations then yield that $I_1 = I_2$. $\square$

7.3 An operation $*$ is called *reducible* if $x*y = x*z$ implies that $y = z$ for all x, y, z. Show that if $(\mathcal{G}, *)$ is a group, then $*$ is reducible. In particular, show that for each $x \in \mathcal{G}$, its inverse is unique.

Solution. From $x*y = x*z$, it follows that $x^{-1}*x*y = x^{-1}*x*z$ and thus $y = z$. (Note that $y = z$ implies $x*y = x*z$, for all x, because in $x*y$, y is replaced by its equal element z. The converse is not true, unless $\mathcal{G}$ is a group and thus x has its inverse. For instance, $x \cdot 0 = y \cdot 0$ does not imply that $x = y$.) The above shows that each inverse in a group is unique, if in $x*y = x*z$ we replace y and z with x^{-1} and $(x^{-1})'$. □

7.5 Let $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ be an isomorphism. Show that $[\widetilde{\mathcal{G}}, \mathcal{G}, f^{-1}]$ is also an isomorphism.

Solution. We only need to show that f^{-1} is a homomorphism from $\widetilde{\mathcal{G}}$ into $\mathcal{G}$. Suppose there are $\widetilde{x}, \widetilde{y} \in \widetilde{\mathcal{G}}$ such that

$$f^{-1}(\widetilde{x} * \widetilde{y}) \neq f^{-1}(\widetilde{x}) * f^{-1}(\widetilde{y}). \tag{P7.5}$$

Then, applying f to the left- and right-hand sides of (P7.5) we get that $\widetilde{x} * \widetilde{y} \neq \widetilde{x} * \widetilde{y}$. □

7.7 Let $[\mathcal{G}, \widetilde{\mathcal{G}}, f]$ be an isomorphism. Find $\mathrm{Ker} f$.

Solution. From Problem 7.6, $\mathrm{Ker} f = I$, because f is bijective. □

7.9 Let $\mathcal{G}$ be the set of all 2×2 real matrices whose determinants equal 1.

a) Show that $(\mathcal{G}, \cdot)$ (with the usual matrix multiplication) is a group.

b) Let B be any 2×2 nonsingular matrix. Define the map $[\mathcal{G}, \mathcal{G}, f]$ such that $f(A) = B^{-1}AB$. Show that $[\mathcal{G}, \mathcal{G}, f]$ is an automorphism.

Solution. $a)$ $\cdot$ is associative. Then, the 2×2 identity matrix I belongs to $\mathcal{G}$, because $\det I = 1$. For each $A \in \mathcal{G}$, A^{-1} exists. Therefore, $\mathcal{G}$ is a group.

b) $[\mathcal{G}, \mathcal{G}, f]$ is a homomorphism. Indeed, by the definition of f,

$$f(AC) = B^{-1}ACB = (B^{-1}AB)(B^{-1}CB) = f(A)f(B).$$

Furthermore,

$$\det(B^{-1}AB) = \det(B^{-1})\det(B)\det(A) = 1.$$

Thus, $B^{-1}AB \in \mathcal{G}$ and f is an endomorphism. Now,

$$f^{-1}(A) = (B^{-1}AB)^{-1} = B^{-1}A^{-1}B.$$

Because $A^{-1} \in \mathcal{G}$ and thus obviously, $\det[f^{-1}(A)]$ equals 1, $f^{-1}(A) \in \mathcal{G}$. Consequently, f is bijective and $[\mathcal{G}, \mathcal{G}, f]$ is an isomorphism. Finally, it all makes $[\mathcal{G}, \mathcal{G}, f]$ an automorphism. □

7.11 Show that l^p (defined in Example 7.11 (v)) is a vector space over $\mathbb{C}$; specifically, show that $x, y \in l^p \Rightarrow x + y \in l^p$. [*Hint*: Apply the inequality in Problem 7.10 in the form $|x_n + y_n|^p \leq 2^{p-1}(|x_n|^p + |y_n|^p)$.]

Solution. First observe that for all $p \geq 1$, from Problem 7.10,

$$|x_n + y_n|^p \leq ||x_n| + |y_n||^p \leq 2^{p-1}(|x_n|^p + |y_n|^p).$$

Thus,

$$\sum_{n=1}^{\infty}|x_n + y_n|^p \leq 2^{p-1}\sum_{n=1}^{\infty}(|x_n|^p + |y_n|^p)$$

$$\leq 2^{p-1}(\sum_{n=1}^{\infty}|x_n|^p + \sum_{n=1}^{\infty}|y_n|^p) < \infty,$$

if $\boldsymbol{x} = \{x_n\}$, $\boldsymbol{y} = \{y_n\} \in l^p$. Hence, $\{x_n + y_n\} \in l^p$. That $\alpha \cdot \boldsymbol{x} \in l^p$ (where $\alpha \in \mathbb{C}$) is trivial. □

7.14 Let S be a subset of $\mathbb{C}$. Argue for what cases S is a subspace of $\mathbb{C}$ over $\mathbb{R}$:

a) S is a closed unit disc centered at zero; that is, $S = \{z \in \mathbb{C} : |z| \leq 1\}$.

b) $S = \{z \in \mathbb{C}\colon \{|\text{Re}(z)| \leq 1\} \times \{|\text{Im}(z)| \leq 1\}\}$.

c) $S = \{z \in \mathbb{C}\colon \{\text{Im}(z) = 0\} \times \{|\text{Im}(z)| \leq 1\}\}$.

d) $S = \{z \in \mathbb{C} : \text{Im}(z) \geq 0 \text{ and } \text{Re}(z) \geq 0\} \cup \{z \in \mathbb{C} : \text{Im}(z) \leq 0 \text{ and } \text{Re}(z) \leq 0\}$.

Solution. S's in *a* - *d*) are not closed relative to the addition and in *a* - *c*) not closed relative to the multiplication of vectors by scalars. Thus, none of S in *a* - *d*) is a subspace. □

7.16 Let $(X, \mathbb{R}, +, \cdot, \preceq)$ be a Riesz space and let $A, B \subseteq X$ such that $\inf A, \inf B, \inf(A + B) \in X$. Prove the following equations.

(i) $\inf(A + B) = \inf A + \inf B$.

Solution. $-\inf(A + B) = \sup(-A - B)$

$= \sup(-A) + \sup(-B) = -\inf A - \inf B$. □

(ii) For any $\alpha \geq 0$, $\alpha \inf A = \inf(\alpha A)$.

Solution. $\alpha \inf A = \alpha(-1)\sup(-A)$

$= -\sup((-1)\alpha A) = \inf(\alpha A)$. □

(iii) For any $\alpha \leq 0$, the following hold: $\alpha \sup A = \inf(\alpha A)$ and $\alpha \inf A = \sup(\alpha A)$.

Solution. Similar to (ii). □

7.18 Let $(X, \mathbb{R}, +, \cdot, \preceq)$ be a Riesz space. Prove the following inequalities.

(i) We say that $A \preceq u$ if $u \in \mathcal{U}(A)$. Let $A \preceq u$ and $B \preceq u$, then $\sup(A \cup B) \preceq u$.

Solution. Obviously, $A \cup B \preceq u$ and thus $\sup(A \cup B) \preceq u$. □

(ii) If $x \preceq z$ and $y \preceq z$, then $x \vee y \preceq z$.

Solution. By (i), $\{x, y\} \preceq z$, because z is an upper bound for sets $\{x\}$ and $\{y\}$, we have $x \vee y \preceq z$. □

(iii) If $x \preceq z$ and $-x \preceq z$, then $|x| \preceq z$.

Solution. Follows directly from (ii) and the definition of $|x|$. □

(iv) The inequality $-z \preceq x \preceq z$ is equivalent to $|x| \preceq z$.

Solution. From $-x \preceq z$ we have by Axiom 2,

$$-x + (x - z) \preceq z + (x - z) \Leftrightarrow -z \preceq x.$$

Thus, $x \preceq z$ and $-x \preceq z$ is equivalent to $-z \preceq x \preceq z$. On the other hand, from $|x| = x \vee (-x) \preceq z$ it follows that

$$x \preceq |x|, -x \preceq |x|, \text{ and } x \preceq z \text{ and } -x \preceq z,$$

which yields $|x| \preceq z$ by (iii). □

(v) The triangle inequality: $|x + y| \preceq |x| + |y|$.

Solution. First notice that if $a \preceq c$ and $b \preceq d$, then $a + b \preceq c + d$. The proof is simple: $a + b \preceq c + b \preceq c + b + (d - b)$. Now, because $a \preceq |a| \Rightarrow x + y \preceq |x| + |y|$ and $-x - y \preceq |x| + |y|$, the statement then follows from (vi). □

8.1 Prove the statement in Proposition 8.8(i) that *the subset* $\mathrm{Ker}f \subseteq \mathcal{V}$ *is a linear manifold.*

Solution. First, by Problem 7.6, $\theta_{\mathcal{V}} \in \mathrm{Ker}f$. Secondly, if $x, y \in \mathrm{Ker}f$, $f(x) = f(y) = \theta_{\mathcal{W}}$, which yields

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha\theta_{\mathcal{W}} + \beta\theta_{\mathcal{W}} = \theta_{\mathcal{W}}$$

[by 7.6 (iv) properties of vector spaces]

implying that $\alpha x + \beta y \in \mathrm{Ker}f$ (by the definition of $\mathrm{Ker}f$ holding for homeomorphisms). Notice that in the proof it was essential that both $\mathcal{V}$ and $\mathcal{W}$ were vector spaces over the same field. □

8.3 Let X be a finite-dimensional vector space and let $Y \subseteq X$ be a linear manifold. Show that $\dim Y \leq \dim X$.

Solution. Let A be a Hamel basis for Y. Thus, A is a linearly independent subset of X. By Proposition 8.13, there is a Hamel basis H for X such that $A \subseteq H$. Thus $\mathrm{card}(A) \leq \mathrm{card}(H)$. □

8.6 Let X be a vector space and let $Y, Z \subseteq X$ be two linear manifolds. Prove the statement. *For each* $x \in Y + Z$*, there is a unique decomposition of* x *by vectors* $y \in Y$ *and* $z \in Z$ *such that* $x = y + z$ *if and only if* $Y \cap Z = \{\theta\}$. (Two such manifolds are said to be *linearly disjoint.*)

Solution. (i) Suppose Y and Z are linearly disjoint. Let $x \in Y + Z$ such that $x = y + z = y' + z'$ with $y, y' \in Y$ and $z, z' \in Z$. Then,

$$Y \ni y - y' = z' - z \in Z$$

implying that $y - y'$ and $z' - z$ belong to either Y and Z. But due to our assumption, Y and Z are linearly disjoint, thus we have $y = y'$ and $z = z'$ proving that indeed $x = y + z$ is a unique decomposition of x.

(ii) Suppose each $x \in Y + Z$ admits a unique decomposition by vectors from Y and Z but suppose that Y and Z are not linearly disjoint. Then there is a vector $w \in Y \cap Z$ such that $w \neq \theta$. By the assumption, $X \ni x = y + z = y - \alpha w + z - \alpha w$, good for all $\alpha \in \mathbb{F}$. But $y - \alpha w \in Y$ and $z - \alpha w \in Z$ meaning that the above decomposition is not unique, which is a contradiction. □

8.7 Let X be a finite-dimensional vector space and $Y_1, Y_2 \subseteq X$ be two linearly disjoint manifolds. (See Problem 8.6.) Then, $Y_1 + Y_2$ is called a *direct sum* of Y_1 and Y_2 and it is denoted by $Y_1 \oplus Y_2$. Prove that if $X = Y_1 \oplus Y_2$, then $\dim X = \dim Y_1 + \dim Y_2$.

Solution. By Problem 8.3, $\dim Y_1 \leq \dim X$ and $\dim Y_2 \leq \dim X$. Thus there are Hamel bases, $H_1 = \{a_1, \ldots, a_k\}$ for Y_1 and $H_2 = \{b_1, \ldots, b_m\}$ for Y_2. Denote $H = \{a_1, \ldots, a_k, b_1, \ldots, b_m\}$. First, we show that H is linearly independent. To render it we let

$$\sum_{i=1}^{k} \alpha_i a_i + \sum_{j=1}^{m} \beta_j b_j = \theta.$$

Because $\theta \in X$ and $\sum_{i=1}^{k} \alpha_i a_i = \ : y_1 \in Y_1$ and $\sum_{j=1}^{m} \beta_j b_j = \ : y_2 \in Y_2$ we have $\theta = y_1 + y_2$ being a decomposition of θ as an element of X by two elements of Y_1 and Y_2, respectively. Due to the assumption that $Y_1 \oplus Y_2$ is a direct sum. Thus, by Problem 8.6, this decomposition is unique because $\theta = \theta + \theta$ is a legitimate representation of θ. We conclude that $\sum_{i=1}^{k} \alpha_i a_i = \sum_{j=1}^{m} \beta_j b_j = \theta$. Then, because H_1 and H_2 are Hamel bases, we have $\alpha_i = \beta_j = 0$ for all relevant i's and j's. Consequently, H is linearly independent.

Secondly, we show that $\operatorname{span} H = X$. Because $X = Y_1 + Y_2$, for each $x \in X, x = y_1 + y_2$ holds, where $y_1 = \sum_{i=1}^{k} \alpha_i a_i$ and $y_2 = \sum_{j=1}^{m} \beta_j b_j$. These imply that $x = \sum_{i=1}^{k} \alpha_i a_i + \sum_{j=1}^{m} \beta_j b_j$ and thus serves as an evidence that $\operatorname{span} H = X$. The latter is equivalent to $\dim X = k + m = \dim Y_1 + \dim Y_2$. □

Chapter 2. Analysis of Metric Spaces

1.2 Let $X = \mathbb{R}$ and $d(x, y) = \sqrt{|x - y|}$. Is (X, d) a metric space?

Solution. d is a metric. The first two properties are evident. To show the triangle inequality, we use

$$\sqrt{|a| + |b|} \leq \sqrt{|a|} + \sqrt{|b|}$$

and

$$\begin{aligned} d(x, y) &= \sqrt{|x - y|} = \sqrt{|x - z + z - y|} \\ &\leq \sqrt{|x - z| + |z - y|} \end{aligned}$$

identifying $x - z$ as a and $z - y$ as b. □

1.4 Let d be a metric on X. Define

$$\rho(x, y) = \tfrac{d(x,y)}{1 + d(x,y)} \ .$$

Show that ρ is a metric on X.

Solution. We need to show just the triangle inequality because the two other properties of ρ as a metric are obvious. Let

$$d_1 = d(x, y), \ d_2 = d(x, z), \text{ and } d_3 = d(z, y).$$

Because $d_1 \leq d_2 + d_3,$ we have that

$$d_1 \leq d_2 + d_3 \leq d_2 + 2d_2d_3 + d_3,$$

which implies that

$$d_1 \left[1 - \tfrac{d_2}{1+d_2} - \tfrac{d_3}{1+d_3}\right] \leq \tfrac{d_2}{1+d_2} + \tfrac{d_3}{1+d_3}.$$

The latter, after division by $1 + d_1,$ is equivalent to

$$d_1 \leq (1+d_1)\tfrac{d_2}{1+d_2} + (1+d_1)\tfrac{d_3}{1+d_3}$$

which yields the desired triangle inequality. □

1.6 (a) Prove Hölder's inequality for finite sums: for conjugate exponents $p > 1$ and $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, $a_1, \ldots, a_n \geq 0,$ and $b_1, \ldots, b_n \geq 0$,

$$\sum_{i=1}^{n} a_i b_i \leq \left[\sum_{i=1}^{n} a_i^p\right]^{1/p} \left[\sum_{i=1}^{n} b_i^q\right]^{1/q}.$$

Hint : Apply Problem 1.5 to $x = a_i/A$ and $y = b_i/B,$ where

$$A = \left[\sum_{i=1}^{n} a_i^p\right]^{1/p} \quad \text{and} \quad B = \left[\sum_{i=1}^{n} b_i^q\right]^{1/q}.$$

(b) Generalize Hölder's inequality for infinite sums.

Solution. $a)$ With the above notation of A and $B,$ assume that $AB > 0$. Otherwise, the inequality is trivial. Now apply Problem 1.5, with $x = a_i/A$ and $y = b_i/B$ to yield

$$\tfrac{a_i}{A}\,\tfrac{b_i}{B} \leq \tfrac{a_i^p}{pA^p} + \tfrac{b_i^p}{qB^p}.$$

After summation over i we have

$$\textstyle\sum_{i=1}^{n} a_i b_i \leq \left(\frac{1}{p} + \frac{1}{q}\right) AB = AB$$

as the Hölder's inequality.

$b)$ The inequality holds when we increase the right-hand side as

$$\sum_{i=1}^{n} a_i b_i \leq \left[\sum_{i=1}^{\infty} a_i^p\right]^{1/p} \left[\sum_{i=1}^{\infty} b_i^q\right]^{1/q}.$$

Finally, it still holds when passing to the limit $n \to \infty$ on the left:

$$\sum_{i=1}^{\infty} a_i b_i \le \left[\sum_{i=1}^{\infty} a_i^p\right]^{1/p} \left[\sum_{i=1}^{\infty} b_i^q\right]^{1/q}. \qquad \square$$

1.9 In Problem 1.8 we defined the Euclidean metric on $\mathbb{R}^n$ by equation (P1.8). This metric can be regarded as

$$d_P(x,y) = \sqrt{\sum_{k=1}^{n} d_k(x_k, y_k)^2}\,, \tag{P1.9}$$

$$x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n), \tag{P1.9a}$$

where $d_k(x_k, y_k)$ is the one-dimensional Euclidean metric on the kth coordinate axis (kth factor space). We can extend this notion and define a metric on the n-times Cartesian product set $Y = Y_1 \times \cdots \times Y_n$ by formulas (P1.9 - P1.9a). This problem asserts that such d_p is indeed a metric on Y. We call this metric the *product metric* and the corresponding metric space (Y, d_P) the product space. In notation, $\times \{(Y_k, d_k) : k = 1, \ldots, n\}$.

Prove the statement: *Let* $(Y_k, d_k),\ k = 1, \ldots, n,$ *be a collection of metric spaces and let* Y *be the Cartesian product of* $Y_1, \ldots, Y_n$. *Then the function* d_P *on* $Y \times Y$ *defined by* (P1.9 - P1.9a) *is a metric on* Y.

Solution. It remains to prove the triangle inequality. Let

$$a_k = d_k(x_k, y_k) \text{ and } b_k = d_k(y_k, z_k),\ k = 1, \ldots, n.$$

Then,

$$d_P(x,y) + d_P(y,z) = \sqrt{\textstyle\sum_{k=1}^{n} a_k^2} + \sqrt{\textstyle\sum_{k=1}^{n} b_k^2}$$

(by Minkowski's inequality)

$$\geq \sqrt{\sum_{k=1}^{n}(a_k + b_k)^2} = \sqrt{\sum_{k=1}^{n}(d_k(x_k, y_k) + d_k(y_k, z_k))^2}$$

(because d_k is also a metric on Y_k)

$$\geq \sqrt{\sum_{k=1}^{n}[d_k(x_k, z_k)]^2} = d_P(x, z). \qquad \square$$

2.1 Show that $\partial A = \partial A^c = \overline{A} \cap \overline{A^c}$.

Solution. By the definition, a point $x \in (X, d)$ is called a boundary point of A if every open ball at x contains points from A and from A^c. On the other hand, a point $x \in X$ is called a closure point of $A \subseteq X$ if every open ball centered at x contains at least one element of A. $\square$

2.3 Show that $[\overline{A^c}]^c \subseteq \overline{A}$.

Solution. We show the equivalent inclusion that $(\overline{A})^c \subseteq \overline{A^c}$. Let $x \in (\overline{A})^c$. Then, $x \notin \overline{A}$ and $x \notin A$, which implies that $x \in A^c$ and thus, $x \in \overline{A^c}$. $\square$

2.5 If $x \in \partial A$, must x be an accumulation point?

Solution. No; see for instance Example 2.12 (v) where 2 is such a point:

(v) Let $A = [0, 1) \cup \{2\}$. Then, $\mathring{A} = (0, 1)$, $\overline{A} = [0, 1] \cup \{2\}$, $A' = [0, 1]$, $\partial A = \{0, 1, 2\}$ (since $A^c = (-\infty, 0) \cup [1, 2) \cup (2, \infty)$, $\overline{A^c} = (-\infty, 0] \cup [1, \infty)$, and $\overline{A} \cap \overline{A^c} = \{0, 1, 2\}$). $\square$

2.7 Let $A \subseteq (X, d)$, where X is an infinite set. Show that if x is an accumulation point of A, then every open set containing x contains infinitely many points of A.

Solution. Suppose there is an open set O containing x and just finitely many points $x_1, \ldots, x_n$ of A, none of which is x. Clearly,

$\varepsilon := \min\{d(x, x_1), \ldots, d(x, x_n)\} > 0$. Then $B(x, \varepsilon)$ contains at most one point of A, which is x if $x \in A$. Consequently, $B(x, \varepsilon) \cap (A \backslash \{x\}) = \emptyset$, that is, x is not an accumulation point of A and this is a contradiction. □

2.8 Give an example of a continuum closed set that does not have any accumulation point.

Solution. Consider the set $A = (a, b)$ $(a < b)$ in the discrete metric space with the carrier $\mathbb{R}$. Then, A does not have any accumulation point, because for each $x \in A$, $B(x, \frac{1}{2}) = \{x\}$. In other words, for each $x \in A$, there is an open ball centered at x containing only x. (Recall that A is closed.) □

2.11 Prove that $\overline{A} = \mathring{A} + \partial A$.

Solution. a) We prove that $\mathring{A} \cap \overline{A^c} = \emptyset$. Let $x \in \mathring{A}$. Then, there is an r such that $B(x, r) \subseteq A \Rightarrow B(x, r) \cap A^c = \emptyset \Rightarrow x \notin \overline{A^c}$.

b) We show that $\mathring{A} + \overline{A^c} = X$ (the carrier), which would be equivalent to $\mathring{A} = \left(\overline{A^c}\right)^c$. Obviously,

$$x \notin \overline{A^c} \Leftrightarrow \exists r\colon B(x, r) \cap A^c = \emptyset \Leftrightarrow B(x, r) \subseteq A \Leftrightarrow x \in \mathring{A}.$$

c) Prove the statement.

$$\begin{aligned} \overline{A} &= \overline{A} \cap X = \overline{A} \cap (\mathring{A} + \overline{A^c}) \\ &= \mathring{A} + (\overline{A} \cap \overline{A^c}) = \mathring{A} + \partial A. \end{aligned}$$

□

2.12 Show that $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

Solution. A closure point of A is obviously a closure point of $A \cup B$. Thus $\overline{A} \subseteq \overline{A \cup B}$ and so is $\overline{B} \subseteq \overline{A \cup B}$. Therefore, $\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$ and we need to show the inverse inclusion. Suppose there is a point $x \in \overline{A \cup B}$ that belongs neither to $\overline{A}$ nor $\overline{B}$. The latter is equivalent to the existence of balls $B(x, r_1)$ and $B(x, r_2)$ that do not meet A

and B, respectively. Thus the ball $B(x,r)$, where $r = \min\{r_1, r_2\}$ does not meet $A \cup B$. But this is absurd, in as much as x is a closure point of $A \cup B$. □

2.15 Prove Corollary 2.14. *A set A is closed if and only if $A = \overline{A}$.*

Solution. Let A be closed. Then A is a closed superset containing A. By Proposition 2.11, $\overline{A}$ is the smallest closed set containing A. Thus, $\overline{A} \subseteq A \Rightarrow A = \overline{A}$. Now, let $A = \overline{A}$. Because $\overline{A}$ is closed, by Proposition 2.11, A is closed. □

2.17 Define $S(x,r) = \{y \in X : d(x,y) = r\}$ and call it a *sphere centered at x with radius r ($r > 0$)*. In Euclidean space $(\mathbb{R}^n, d_e)$, obviously, $S(x,r) = \partial B(x,r)$. Is this true for metric spaces? If it is not, then is there at least a relationship between $S(x,r)$ and $\partial B(x,r)$ common for all metric spaces? Are there any exceptional cases? Explain.

Solution. (i) Due to Remark 2.15,

$$B(x,r) \subseteq \overline{B(x,r)} \subseteq C(x,r).$$

The inequalities will continue to hold if we intersect each of the three sets above with one and the same set. Therefore, the inequalities will hold if we subtract a set yielding

$$\emptyset = B(x,r) \backslash B(x,r) \subseteq \overline{B(x,r)} \backslash B(x,r)$$

$$\subseteq C(x,r) \backslash B(x,r). \tag{P2.17}$$

(ii) Now, obviously, $S(x,r) = C(x,r) \backslash B(x,r)$. (P2.17a)

(iii) On the other hand, by Problem 2.1,

$$\partial B(x,r) = \overline{B(x,r)} \cap \overline{(B(x,r))^c} = \overline{B(x,r)} \cap (B(x,r))^c$$

$$= \overline{B(x,r)} \backslash B(x,r), \tag{P2.17b}$$

because $(B(x,r))^c$ is closed. From (P2.17-P2.17b) we thus have

$$\partial B(x,r) \subseteq S(x,r)$$

holding for all metric spaces.

(iv) However, $\partial B(x,r) \neq S(x,r)$, as we can see it from the example of the discrete metric, in which $\partial B(x,r)$ is a proper subset of $S(x,r)$. Indeed, from Example 2.2 (iv), Remark 2.15, and (P2.17a),

$$S(x,r) = \begin{cases} \emptyset, & r \neq 1 \\ \{x\}^c, & r = 1 \end{cases}$$

whereas by Problem 2.16, $\partial B(x,r) = \emptyset$ for all r. □

3.2 Prove that a set $A \subseteq (X,d)$ is nowhere dense in X if and only if the complement of its closure is dense in X.

Solution. Let $B = \overline{A}$. By problem 2.11 part (b) (in its solution), $\overset{\circ}{B} = \emptyset \Leftrightarrow \overline{B^c} = X$.

An alternative and longer proof. We prove that "A is nowhere dense" is equivalent to "$(\overline{A})^c$ is dense in X." Let B be a subset of X such that $\overline{B^c} = X$ and suppose that $\overset{\circ}{B} \neq \emptyset$. Then, there is an $x \in \overset{\circ}{B} \subseteq B$. Then, there is a sequence $\{x_n\} \subseteq B^c$ such that $x_n \to x$. Because x is an interior point of B, there is an open ball $B(x,r) \subseteq B$. Furthermore, there is an r-tail $\{x_k, \dots\}$ (of $\{x_n\}$) entirely in $B(x,r)$, thus in B, not in B^c, which is a contradiction. Thus, $\overset{\circ}{B} = \emptyset$. Conversely, let $B \neq \emptyset$ (because the proof would be trivial), but $\overset{\circ}{B} = \emptyset$. Then, for each $x \in B$ and for each $r > 0$, $B(x,r) \cap B^c \neq \emptyset$. Consider the monotone decreasing sequence $B(x, \frac{1}{n})$ of balls from each of which we select a point $x_n \in B^c$. [Such a point can be chosen because of $B(x,r) \cap B^c \neq \emptyset$.] Clearly, $x_n \to x$. If $x \in B^c$, the same statement is trivial. Thus, $\overline{B^c} = X$. Finally, we take $\overline{A}$ for B. □

3.3 Assuming that $(\mathbb{R}, d_e)$ is complete (a known fact from calculus) prove that $(\mathbb{R}^n, d_e)$ is also complete.

Solution. Let $\{x^{(k)}\}$ be a Cauchy sequence in $(\mathbb{R}^n, d_e)$. Denote $\{x_i^{(k)}\}$ the ith component sequence of $\{x^{(k)}\}$. Because

$$d_e(x_i^{(k)}, x_i^{(m)}) \leq d_e(x^{(k)}, x^{(m)}),$$

each component sequence of $\{x^{(k)}\}$ is also Cauchy and thus convergent in $(\mathbb{R}, d_e)$. Let $x_i^{(k)} \to x_i$ and $x = (x_1, \ldots, x_n)$. Then, for each $\varepsilon/\sqrt{n} > 0$, there is an N_i such that for each $k \geq N_i$,

$$\mid x_i^{(k)} - x_i \mid \; < \varepsilon/\sqrt{n}.$$

Let $N = \max\{N_1, \ldots, N_n\}$. Then, for each $k \geq N$,

$$d_e(x^{(k)}, x) = \sqrt{\textstyle\sum_{i=1}^n (x_i^{(k)} - x_i)^2} < \varepsilon. \qquad \square$$

3.4 Show that any Cauchy sequence is bounded.

Solution. Recall that a sequence $\{x_n\}$ is called bounded if for every n, $d(x_1, x_n) \leq M$, where M is a positive real number. Now, if $\{x_n\}$ is Cauchy, for every $\varepsilon > 0$ there an $N(\varepsilon)$-tail of $\{x_n\}$. It means that $d(x_m, x_n) < \varepsilon$ if $m, n \geq N$. Let $D := \max\{d(x_1, x_s); s = 1, \ldots, N\}$. Then,

$$d(x_1, x_n) \leq d(x_1, x_N) + d(x_N, x_n) \leq D + \varepsilon,$$

for all $n \geq N$, and

$$d(x_1, x_n) \leq D, \text{ for all } n \leq N. \qquad \square$$

3.7 Show that if $\{x_n\} \subseteq (X, d)$ is a Cauchy sequence and $\{x_{n_k}\}$ is a subsequence convergent to a point $a \in X$, then $x_n \to a$.

Solution. $d(x_n, a) \leq d(x_n, x_{n_k}) + d(x_{n_k}, a)$. Now, because $\{x_n\}$ is Cauchy, for each $\varepsilon > 0$, there is an N_1-tail of $\{x_n\}$, and thus of $\{x_{n_k}\}$, such that $d(x_n, x_{n_k}) < \varepsilon$ for all terms of this tail, and $d(x_{n_k}, a) < \varepsilon$, for all $n_k \geq N_2$, due to the convergence of $\{x_{n_k}\}$. Therefore, $d(x_n, a) < 2\varepsilon$, for all $n \geq N = \max\{N_1, N_2\}$. □

3.8 Show that in $(\mathbb{R}^n, d_e)$, $\overline{B(x,r)} = C(x,r)$, where $C(x,r)$ is a closed ball of Remark 2.15.

Solution. a) Because $C(x,r)$ is a closed set and because $B(x,r) \subseteq C(x,r)$, $\overline{B(x,r)} \subseteq C(x,r)$ (which is also due to Remark 2.15).

b) Show the inverse inclusion. Let $y \in C(x,r)$. Define $y_n = \frac{1}{n}x + (1 - \frac{1}{n})\, y$. From

$$d(x, y_n) = \sqrt{\textstyle\sum_{k=1}^{n} [x_k - \frac{1}{n}x_k - (1 - \frac{1}{n})\, y_k]^2}$$

$$\leq \left(1 - \tfrac{1}{n}\right) r < r,$$

it follows that $y_n \in B(x,r)$. Furthermore, $d_e(y_n, y) \to 0$. Thus, by Theorem 3.3 (*A point x is a closure point of a set A if and only if x is a sequential limit point of A.*), $y \in \overline{B(x,r)}$. □

4.1 Show that if A is totally bounded, then A is bounded. Give an example, where a bounded set is not totally bounded.

Solution. (i) Let A be totally bounded. Then, $\forall \varepsilon > 0, \exists$ a finite cover of A by balls of radius ε: $\bigcup_{i=1}^{m} B(x_i, \varepsilon) \supseteq A$. Let $x, y \in A$. Then, $x \in B(x_i, \varepsilon)$ and $y \in B(x_j, \varepsilon)$ implying that

$$d(x,y) \leq d(x, x_i) + d(x_i, x_j) + d(x_j, y)$$

$$< \varepsilon + M + \varepsilon = 2\varepsilon + M,$$

where $M :\, = \max\{d(x_p, x_r) : p, r = 1, \ldots, m\}$. Consequently,

$$d(A) = \sup\{d(x,y) : x, y \in A\} \leq 2\varepsilon + M$$

making A bounded.

(ii) Consider $(\mathbb{R}, d)$, with $d(x,y) = \min\{|x-y|, 1\}$ and $A = [0, \infty)$. Then, A is bounded, but obviously not totally bounded. □

4.3 Show that a continuous bounded function on a bounded interval need not be uniformly continuous.

Solution. Let $f(x) = \sin\frac{1}{x}$ on $I = (0, \frac{2}{\pi}]$. Clearly, f is continuous on I. Let

$$x_n = \tfrac{2}{(2n+1)\pi},\ y_n = \tfrac{1}{n\pi},\ n = 1, 2, \dots\ .$$

Then $f(x_n) = \pm 1$ and $f(y_n) = 0$. Therefore, $|f(x_n) - f(y_n)| = 1$. On the other hand,

$$|x_n - y_n| = \tfrac{1}{n(2n+1)\pi},$$

which can be made arbitrarily small. Therefore, for any $\varepsilon < 1$, there is no δ good for all $x \in I$ required by uniform continuity.

Another solution. Take a sequence $\{x_n = \frac{2}{\pi n}\} \subseteq I$, which is clearly Cauchy. However, $f(x_n) = \sin(\frac{\pi n}{2})$ is an oscillating, non-Cauchy sequence. Thus f is not Cauchy continuous and by Proposition 4.16 f is not uniformly continuous. □

In Problems 4.4 - 4.8 it is assumed that f and g are functions from $(\mathbb{R}, d_e)$ to $(\mathbb{R}, d_e)$.

4.5 Let $f : A \to \mathbb{R}$ be a differentiable function such that its derivative f' is bounded over A, where A is an arbitrary (bounded or unbounded) interval. Show that f is uniformly continuous on A.

Solution. From the mean value theorem,

$$f(x) - f(y) = f'(c)(x - y), \text{ for a } c \in (x, y).$$

Hence, $|f(x) - f(y)| \leq M|x - y|$. For an $\varepsilon > 0$, take $\delta = \frac{\varepsilon}{M}$. Then, $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$. □

4.6 Show that if f and g are uniformly continuous on $\mathbb{R}$ and bounded then fg is uniformly continuous on $\mathbb{R}$ too. However, if f and g are just uniformly continuous then fg need not be uniformly continuous.

Solution. (i) Let $|f| \leq M_1$ and $|g| \leq M_2$ and $M = \max\{M_1, M_2\}$. Select $\varepsilon > 0$ and let $\delta > 0$ be such that $\forall x, y \in X$ with $|x - y| < \delta$, $|f(x) - f(y)| < \frac{\varepsilon}{2M}$ and $|g(x) - g(y)| < \frac{\varepsilon}{2M}$. Then,

$$|f(x)g(x) - f(y)g(y)| \leq |f(x)||g(x) - g(y)|$$

$$+ |g(y)||f(x) - f(y)| < M\frac{\varepsilon}{2M} + M\frac{\varepsilon}{2M} = \varepsilon.$$

(ii) Although the function $f(x) = x$ is uniformly continuous, $g(x) = x^2 = x \cdot x$ is not. □

4.8 Let f be a continuous function and g a uniformly continuous function on a set A such that $|f| \leq |g|$. Is f then uniformly continuous?

Solution. No. For instance, if $f(x) = \sin\frac{1}{x}$ on $I = (0, \frac{2}{\pi}]$ (of Problem 4.3), then $|f| \leq \mathbf{1}$. □

4.10 Show that in $\mathbb{R}^n$, Euclidean and supremum metrics are equivalent.

Solution. The statement follows from two inequalities,

$$d_e(x, y) \leq \sqrt{n}d_\infty(x, y), \quad d_\infty(x, y) \leq d_e(x, y),$$

and Remark 4.9. □

4.12 Let (X, d) be a metric space and define the function $\rho : (X^2, X^2) \to \mathbb{R}_+$ as $\rho(\mathbf{x}, \mathbf{y}) = d(x_1, y_1) + d(x_2, y_2)$, where $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$. (a) Show that ρ is a metric on X^2. (b) Define the function $f(\mathbf{x}) = d(x_1, x_2)$. Show that $f : X^2 \to \mathbb{R}_+$ is uniformly continuous on (X^2, ρ).

Solution. (a) is obvious.

$$(b) \qquad |f(\mathbf{x}) - f(\mathbf{y})| = |d(x_1, x_2) - d(y_1, y_2)|$$

$$= |d(x_1, x_2) - d(y_1, x_2) + d(x_2, y_1) - d(y_1, y_2)|$$

$$\leq |d(x_1, x_2) - d(y_1, x_2)| + |d(x_2, y_1) - d(y_1, y_2)|$$

$$\leq d(x_1, y_1) + d(x_2, y_2) = \rho(\mathbf{x}, \mathbf{y}). \qquad \square$$

4.14 Give an example of a continuous function $[X, Y, f]$ from (X, d) to (Y, ρ) such that f is not Cauchy continuous. (*Hint*: See Example 4.15.)

Solution. The function $f(x) = \arctan x$ is convergent to $\frac{\pi}{2}$ when $x \to \infty$. Thus so does the sequence $x_n = \arctan(n) \to \frac{\pi}{2}$ when $n \to \infty$. Therefore, $\{x_n\}$ must also be Cauchy. However, $\{\text{rtan}(x_n)\} = \{n\}$ is clearly not Cauchy. □

4.16 Prove the following statement. *Let* (X, d) *and* (Y, ρ) *be two metric spaces and let* $[X, Y, f]$ *and* $[X, Y, g]$ *be two continuous mappings. The subset* $A = \{x \in X : f(x) = g(x)\}$ *is closed.*

Solution. A is closed if and only if it contains all of its sequential limit points. Let $x \in \overline{A}$ and let $\{x_n\} \subseteq A$ such that $x_n \to x$. Thus, for each n, $F(x_n) : = f(x_n) - g(x_n) = 0$. Because $F = f - g$ is continuous, $F(x_n) \to F(x)$, because $\{x_n\}$ is convergent. But

obviously $F(x_n) \to 0$ implying that $f(x) = g(x)$ and thus $x \in A \Rightarrow \overline{A} \subseteq A$.

Another Solution. We show that A^c is open. Let $a \in A^c$ and $\rho(f(a), g(a)) = 3M$. Because $a \in A^c$, $M > 0$. Because f and g are continuous, for each $\varepsilon > 0$, there is a $\delta(a) > 0$ such that

$$\rho(f(x), f(a)) < \varepsilon \text{ and } \rho(g(x), g(a)) < \varepsilon$$

for all x such that $d(x, a) < \delta(a)$. Therefore,

$$\begin{aligned} 3M &= \rho(f(a), g(a)) \\ &\leq \rho(f(x), f(a)) + \rho(g(x), g(a)) + \rho(f(x), g(x)) \\ &< 2\varepsilon + \rho(f(x), g(x)). \end{aligned}$$

Let now $\varepsilon = M$. Then for all x such that $d(x, a) < \delta(a)$, we have that $\rho(f(x), g(x)) > M$, that is, for all $x \in B_d(a, \delta(a))$, $f(x) \neq g(x)$. Therefore, $B_d(a, \delta(a)) \subseteq A^c$ yielding that each point $a \in A^c$ is interior. □

5.1 Using similar arguments as in Example 5.6, show that the limit of any uniformly convergent sequence of continuous bounded functions from (X, d_0) to (Y, d) is a bounded and continuous function.

Solution. Let $\{f_n\} \subseteq \mathcal{F}_*((X,d_0),(Y,d))$ be a sequence of continuous bounded functions. Suppose that f is the limit of $\{f_n\}$ in the sense of the supremum metric ρ. Thus, for each $\varepsilon > 0$, say $\varepsilon = 1$, there is an $N = N(1)$ such that $\rho(f_n,f) < \varepsilon$ for all $n > N$. In particular, $f_N \in \mathcal{F}_*$, therefore

$$\rho(f, \mathbf{0}) \leq \rho(f, f_N) + \rho(f_N, \mathbf{0}) < 1 + M(f_N) : = M.$$

Hence $f \in \mathcal{F}_*$. Now we show that f is continuous. Given an x and n, for each $\varepsilon > 0$, there is a $\delta_x > 0$ such that for every y with

$d_0(x,y) < \delta_x,\ d(f_n(x), f_n(y)) < \varepsilon$. Now, by the triangle inequality used twice,

$$d(f(x), f(y)) \le d(f(x), f_n(x))$$

$$+ d(f_n(x), f_n(y)) + d(f_n(y), f_n(y))$$

$$< \rho(f, f_n) + \varepsilon + \rho(f, f_n).$$

Taking $n \ge N,$ we have the last expression dominated by 3ε. This yields continuity of f. □

5.3 Let $\{C_n\}$ be a sequence of closed balls in $(\mathbb{R}^n, d_e)$ such that each of the balls C_n is centered at a point $x_0 \in \mathbb{R}^n$ and has radius $\frac{1}{n}$, $n = 1, 2, \dots$. Find the intersection $\bigcap_{n=1}^{\infty} C_n$.

Solution. By Cantor's Theorem 5.4, $\bigcap_{n=1}^{\infty} C_n$ has exactly one point. It is easily deduced that x_0 is the one. □

5.6 Let $\mathcal{F} \subseteq \mathcal{F}_*(X; (Y, \rho)),$ where Y is a vector space. Prove that $\mathcal{F}$ is d_∞-bounded if and only if there is a positive constant M such that for all $f \in \mathcal{F},\ d_\infty(f, \mathbf{0}) \le M$.

Solution. Recall [Example 5.6(i)] that $\mathcal{F}$ is ρ-bounded or uniformly bounded if $\operatorname{diam} \mathcal{F} = \sup\{d_\infty(f,g) : f, g \in \mathcal{F}\} < \infty$.

$a)$ Suppose $\mathcal{F}$ is such that for each $f \in \mathcal{F},\ d_\infty(f, \mathbf{0}) \le M$. Then, for all $f, g \in \mathcal{F}$,

$$d_\infty(f,g) \le d_\infty(f, \mathbf{0}) + d_\infty(g, \mathbf{0}) \le 2M.$$

Therefore, $\operatorname{diam} \mathcal{F} \le 2M$.

$b)$ Suppose that $\mathcal{F}$ is d_∞-bounded, i.e. $\operatorname{diam} \mathcal{F} < \infty$. We need to show that $d_\infty(f, \mathbf{0}) \le M$ for all $f \in \mathcal{F}$. If $\mathcal{F}$ is a finite subset, then the proof is trivial. Suppose $\mathcal{F}$ is infinite and let $\sup\{d_\infty(f, \mathbf{0})$:

$f \in \mathcal{F}\} = \infty$. Then, there is a sequence $\{f_n\} \subseteq \mathcal{F}$ such that $d_\infty(f_n, \mathbf{0}) \to \infty$. Because d_∞ is a metric and thus $d_\infty(g, \mathbf{0}) < \infty$ for any $g \in \mathcal{F}$, we have

$$d_\infty(f_n, g) \geq |d_\infty(f_n, \mathbf{0}) - d_\infty(g, \mathbf{0})| \to \infty.$$

This contradicts the assumption that diam $\mathcal{F} < \infty$. □

5.8 We can make use of the fact that the Euclidean and uniform (supremum) metrics are equivalent (see Problem 4.10) to show completeness of $(\mathbb{R}^n, d_e)$. For $n = 1$, it is well known from calculus. Prove completeness of $(\mathbb{R}^n, d_e)$ for an arbitrary n. (See Problem 3.5.)

Solution. *a*) We show that d_e and the supremum (uniform) metric are equivalent on $\mathbb{R}^n$. Let $x^{(k)} \overset{d_e}{\to} x$. Thus, for each $\varepsilon > 0$, there is an N such that for all $k \geq N$,

$$d_e(x^{(k)}, x) = \sqrt{\textstyle\sum_{i=1}^n (x_i^{(k)} - x_i)^2} < \varepsilon.$$

Now, obviously,

$$d_\infty(x^{(k)}, x) = \max\left\{ \left| x_i^{(k)} - x_i \right|; i = 1, \ldots, n \right\} \leq d_e(x^{(k)}, x),$$

which yields that $d_\infty(x^{(k)}, x) < \varepsilon$ as well. Conversely, let

$$x^{(k)} \overset{d_\infty}{\to} x.$$

Then, because $d_e(x^{(k)}, x) \leq \sqrt{n}\, d_\infty(x^{(k)}, x)$, we easily arrive at

$$x^{(k)} \overset{d_e}{\to} x.$$

b) Now, let $\{x^{(k)}\}$ be a Cauchy sequence in $(\mathbb{R}^n, d_e)$. Then, $\{x_i^{(k)}\}$, for each $i = 1, \ldots, n$, is Cauchy in $(\mathbb{R}, d_e)$ [because $d_e(x_i^{(k)}, x_i^{(m)}) \leq$

$d_e(x^{(k)}, x^{(m)})]$. But then, for each i, $\{x_i^{(k)}\}$ is convergent, as $(\mathbb{R}, d_e)$ is complete. Let

$$x_i^{(k)} \xrightarrow{d_e} x_i.$$

Thus,

$$\max\{d_e(x_i^{(k)}, x_i);\, i = 1, \ldots, n\} \to 0,$$

i.e.,

$$x^{(k)} \xrightarrow{d_\infty} x.$$

Due to the equivalence of Euclidean and supremum metrics (a),

$$x^{(k)} \xrightarrow{d_e} x. \qquad \square$$

6.1 Show that if $\{x_k\} \subseteq (\mathbb{R}^n, d_e)$ with $d(x_k, 0) \leq 3$, then $\{x_k\}$ has a convergent subsequence.

Solution. Let $A = \{x \in \mathbb{R}^n\colon d(x, 0) \leq 3\}$. Then A, as a closed ball, is closed and bounded, thus compact. By Theorem 6.3, every sequence in A must have a convergent subsequence. $\square$

6.3 Let $A, B \subseteq (X, d)$ such that A is compact and B is closed. If $A \cap B = \emptyset$, show that $d(A, B) > 0$.

Solution. First observe that

$$d(\{x\}, B) \text{ is } \begin{cases} = 0, & x \in \overline{B} \\ > 0, & x \in (\overline{B})^c. \end{cases}$$

Indeed: 1) If $x \in \overline{B}$, for each $\varepsilon > 0$, $B(x, \varepsilon) \cap B \neq \emptyset$. Thus, there is $y_B \in B(x, \varepsilon) \cap B$. Clearly, $d(x, y_B) < \varepsilon$ implying that

$$d(\{x\}, B) \leq d(x, y_B) < \varepsilon$$

$\Rightarrow\ d(\{x\}, B) = 0.$

2) If $x \in (\overline{B})^c$, there is an open ball $B(x,r)$ such that $B(x,r) \cap \overline{B} = \emptyset \quad \Rightarrow \quad \forall y_B \in B, \quad d(x,y_B) > r$. Obviously, the converses to each of these statements hold.

Now, by Problem 6.2, there is an $a \in A$ such that $d(\{a\},B) = d(A,B)$. If $d(A,B) = 0$, then $d(\{a\},B) = 0$ and hence $a \in \overline{B}$. However, because B is assumed to be closed, $a \in B$ and then $A \cap B \neq \emptyset$, which is not the case, by the assumption. Thus,

$$d(A,B) > 0. \qquad \square$$

6.4 Let $A \subseteq (X,d)$. Show that if A is totally bounded then $\overline{A}$ is also totally bounded.

Solution. Choose some $\varepsilon > 0$. Let $x \in \overline{A}$. Then, $B(x, \frac{\varepsilon}{2}) \cap A \neq \emptyset$. Let $y \in B(x, \frac{\varepsilon}{2}) \cap A$. Then $d(x,y) < \frac{\varepsilon}{2}$. Because A is totally bounded,

$$A \subseteq \bigcup_{i=1}^{n} B(x_i, \tfrac{\varepsilon}{2}).$$

Thus, $y \in B(x_i, \frac{\varepsilon}{2})$, for some i and

$$d(x,x_i) \leq d(x,y) + d(y,x_i) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Therefore, $x \in B(x_i,\varepsilon)$. Because $A \subseteq \bigcup_{i=1}^{n} B(x_i,\varepsilon)$ and an arbitrary $x \in \overline{A}$ belongs to one of the balls of the ε-cover of A, it follows that $\{B(x_1,\varepsilon),\ldots,B(x_n,\varepsilon)\}$ also covers $\overline{A}$. $\square$

6.6 Show that sequential compactness of a subspace implies its completeness.

Solution. Let $\{x_n\}$ be a Cauchy sequence in A. Then, $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$. Denote the limit of this subsequence by x. Then, for each $\varepsilon > 0$, there is an N' such that for all $k \geq N'$, $d(x_{n_k},x) < \varepsilon$. On the other hand, because $\{x_n\}$ is Cauchy, there is an N such that for all $m,n \geq N$, $d(x_n,x_m) < \varepsilon$. Now,

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < 2\varepsilon,$$

when $n, n_k \geq \max\{N, N'\}$, which shows that x is also the limit of $\{x_n\}$. □

6.7 Prove Theorem 6.2. *Any subset of $\mathbb{R}^n$ is Lindelöf set in $(\mathbb{R}^n, d_e)$. In particular, $(\mathbb{R}^n, d_e)$ is a Lindelöf space.*

Solution. (i) Let $\{O_i : i \in I\}$ be an open cover of $\mathbb{R}^n$ and let $x \in \mathbb{R}^n$. Then, there is an element O of the cover $\{O_i : i \in I\}$ that contains x. By Problem 2.19, there is an open ball $B(q, r) \subset O$ such that $q \in \mathbb{Q}^n, r \in \mathbb{Q}$, and $x \in B(q, r)$. Consequently, we can cover $\mathbb{R}^n$ by all such balls, each associated with $x \in \mathbb{R}^n$. Because each such ball has rational parameters q and r, the entire set of all these balls is countable. Let us enumerate these balls as $\{B(q_k, r_k) : k \in \mathbb{N}\}$. Obviously, each of these balls contains more than one points such as x above.

Now, because $B(q_k, r_k)$ is contained by some O, say O_{i_k}, it follows that $\{O_{i_k} : k \in \mathbb{N}\}$ also covers $\mathbb{R}^n$, but in this case $\{O_{i_k} : k \in \mathbb{N}\}$ is a countable reduction of the original cover $\{O_i : i \in I\}$ proving that $(\mathbb{R}^n, d_e)$ is Lindelöf.

(ii) The proof is identical if we replace $\mathbb{R}^n$ with a subset $A \subseteq \mathbb{R}^n$. □

6.9 Show that a mapping $f : (X, d) \to (Y, \rho)$ is continuous if and only if, for any compact subset K of X, $\text{Res}_K f$ is continuous.

Solution. (i) Let $\mathcal{K}$ be the set of all compact subsets of X and f be continuous on each $K \in \mathcal{K}$. In other words, given a $K \in \mathcal{K}$, $\text{Res}_K f$ is continuous. We show that f is continuous on X. Throughout the proof, we will be using Heine's Theorem 4.4: *A function f: $(X, d) \to (Y, \rho)$ is continuous at $x \in X$ if and only if for every sequence $\{x_n\}$, d-convergent to x, its image sequence $\{f(x_n)\}$ is ρ-convergent to $f(x)$.*

Let $\{x_n\}$ be any convergent sequence, say to a point x and let $A = \{x, x_1, x_2, \dots\}$. Obviously, A is sequentially compact. Thus, by Theorem 6.3, $A \in \mathcal{K}$. By the assumption, f is continuous on A. Thus, according to Heine, $f(y_n) \to f(y)$ for every convergent sequence $\{y_n\} \subseteq A$. In particular, it should work for $\{x_n\}$, that is $f(x_n) \to f(x)$. In summary, we showed that given an arbitrary sequence $\{x_n\}$, convergent to a point x, $f(x_n) \to f(x)$. This, again, by Heine, implies that f is continuous on the entire X.

(ii) The converse. Suppose f is continuous on X and let $K \in \mathcal{K}$. Then, by Theorem 6.3, K is sequentially compact. In particular, it means that if $\{x_n\}$ is a sequence convergent to a point x and if $\{x_n\} \subseteq K$, then $x_n \to x$. (See Problem 3.7.) In an alternative argument, by Theorem 6.7, K is closed and bounded. Thus, by Corollary 3.4, K contains all of its sequential limits. Thus, if $\{x_n\}$ is a convergent sequence in K, its limit, say x, must belong to K. Then, by Heine's Theorem, because f is continuous on X, $f(x_n) \to f(x)$. But because $\{x, x_1, x_2, \dots\} \subseteq K$, it follows that f has the Heine's property on K, and because $\{x_n\}$ was an arbitrary convergent sequence in K, with all said above, f must be continuous on K, i.e., $\text{Res}_K f$ is continuous. □

7.2 Define the space l^∞ as the set of all bounded sequences $x = \{x_1, x_2, \dots\} \subseteq \mathbb{C}$. Show that l^∞ is an NVS with the norm defined as $\|x\| = \sup\{|x_i|\colon\ i = 1, 2, \dots\}$.

Solution. We show the triangle inequality. Other properties of the supremum norm are obvious.

$$|x_i + y_i| \le |x_i| + |y_i| \le \|x\| + \|y\|.$$

Thus,

$$\|x + y\| = \sup\{|x_i + y_i|\} \le \|x\| + \|y\|.$$

To show linearity, observe that if $x, y \in l^\infty$, then x and y are bounded and so is $\alpha x + \beta y$. □

7.4 Let $\mathcal{F}_*(\Omega)$ be the space of all bounded real-valued functions on Ω. Show that $\mathcal{F}_*$ is a vector space. Let $\|f\|_\infty = \sup\{|f(\omega)| : \omega \in \Omega\}$ be the supremum norm defined in Example 7.7(iii). Show that the supremum norm in $\mathcal{F}_*$ is indeed a norm and show that $\mathcal{F}_*$ is a Banach space with respect to this norm.

Solution. That $\|\cdot\|_\infty$ is a norm can be readily shown as in Problem 7.2. We show that $\mathcal{F}_*(\Omega)$ is Banach. Let $\{f_n\}$ be a Cauchy sequence in $\mathcal{F}_*(\Omega)$. Then, given $\varepsilon = 1,$ there is $N = N(1)$ such that for all $m, n \geq N(1),\ \|f_n - f_m\|_\infty < 1$. Let $m = N$. Then,

$$\begin{aligned}\|f_n\|_\infty &\leq \|f_n - f_N\|_\infty + \|f_N\|_\infty \\ &< 1 + M_N := M,\ n \geq N,\end{aligned}$$

that is, $\{f_n\}$ is uniformly bounded. The latter is also due to Example 5.6(i). Because the Euclidean norm is complete, by Example 5.6(ii) and Problem 5.7, $(\mathcal{F}_*(\Omega), \|\cdot\|_\infty)$ is Banach. □

7.5 Show that $\|\cdot\|_p$ in Example 7.7(ii) is a norm.

Solution. By Minkowski's inequality for infinite sums (Problem 1.7*b*),

$$\left[\sum_{i=1}^{\infty} |\, x_i + y_i \,|^{\,p}\right]^{1/p} \leq \left[\sum_{i=1}^{\infty} |\, x_i \,|^{\,p}\right]^{1/p} + \left[\sum_{i=1}^{\infty} |\, y_i \,|^{\,p}\right]^{1/p}$$

which is the triangle inequality. □

7.7 Show that the pointwise limit x of the sequence $\{x^{(n)}\}$ in Example 7.7(ii) is also an l^p-limit.

Solution. We show that $x^{(n)} \to x$ in the l^p norm. Because $\{x_k\}$ is Cauchy, for each $\varepsilon > 0,$ there is an $N(\varepsilon)$-tail of $\{x_k\}$. Hence for all $m, n \geq N,$

$$\sum_{k=1}^{\infty}\left|x_k^{(n)}-x_k^{(m)}\right|^p<\varepsilon^p$$

and consequently,

$$\sum_{k=1}^{r}\left|x_k^{(n)}-x_k^{(m)}\right|^p<\varepsilon^p,\ \ r=1,2,\ldots$$

Letting $m\to\infty$ we have from the last equation

$$\sum_{k=1}^{r}\left|x_k^{(n)}-x_k\right|^p\le\varepsilon^p,\ \ r=1,2,\ldots,n\ge N$$

and then with $r\to\infty$ we still have

$$\sum_{k=1}^{\infty}\left|x_k^{(n)}-x_k\right|^p\le\varepsilon^p,\ \ n\ge N.$$

Thus, $x^{(n)}\to x$ in the l^p-norm. □

8.2 Prove that the function $\|\cdot\|_*$ in Definition 8.10 is a norm.

Solution. (i) $x=\theta\ \Leftrightarrow$ all $\alpha_i=0\ \Rightarrow\|x\|_*=0\ \Leftrightarrow x=\theta$.

(ii) $\|\gamma x\|_*=\sqrt{\sum_{i=1}^{n}|\gamma\alpha_i|^2}=|\gamma|\|x\|_*$.

$$\begin{aligned}(iii)\qquad \|x+y\|_*^2&=\|\textstyle\sum_{i=1}^{n}(\alpha_i+\beta_i)b_i\|_*\\&=\textstyle\sum_{i=1}^{n}|\alpha_i+\beta_i|^2\\&=\textstyle\sum_{i=1}^{n}(\alpha_i+\beta_i)(\overline{\alpha}_i+\overline{\beta}_i)=\sum_{i=1}^{n}\Big[|\alpha_i|^2+|\beta_i|^2\\&+2\mathrm{Re}(\alpha_i\beta_i)\Big]\le\textstyle\sum_{i=1}^{n}\Big[|\alpha_i|^2+|\beta_i|^2+2|\alpha_i||\beta_i|\Big]\\&=\textstyle\sum_{i=1}^{n}(|\alpha_i|+|\beta_i|)^2\end{aligned}$$

(by Problem 1.6, Hölder's inequality for finite sums: for conjugate exponents $p = q = \frac{1}{2}$ and $a_1, \ldots, a_n \geq 0$, and $b_1, \ldots, b_n \geq 0$

$$\sum_{i=1}^{n} a_i b_i \leq \left[\sum_{i=1}^{n} a_i^2\right]^{1/2} \left[\sum_{i=1}^{n} b_i^2\right]^{1/2}\Big)$$

$$\leq \sum_{i=1}^{n} \left[|\alpha_i|^2 + |\beta_i|^2\right]$$

$$+ 2\sqrt{\sum_{i=1}^{n} |\alpha_i|^2} \cdot \sqrt{\sum_{i=1}^{n} |\beta_i|^2}$$

$$= (\|x\|_* + \|y\|_*)^2. \qquad \square$$

Problems 8.4-8.6 offer an alternative proof to Theorem 8.9 by showing that any norm $\|\cdot\|$ on X is equivalent to the standard norm $\|\cdot\|_*$.

8.4 Let $(X, \|\cdot\|)$ be a finite-dimensional NVS and $\|\cdot\|_*$ be the standard norm defined in (8.13). Show that there is a positive constant M such that for each $x \in X$,

$$\|x\| \leq M\|x\|_*.$$

Solution. Let $\{b_1, \ldots, b_n\}$ be a Hamel basis for X. Define

$$M = \sqrt{\sum_{i=1}^{n} \|b_i\|^2} > 0.$$

Then if $\|\cdot\|$ is a norm on X,

$$\|x\| = \|\textstyle\sum_{i=1}^{n} \alpha_i b_i\| \leq \sum_{i=1}^{n} |\alpha_i| \|b_i\|$$

and by Hölder's inequality (Problem 1.6)

$$\leq M\sqrt{\sum_{i=1}^{n} |\alpha_i|^2} = M\|x\|_*. \qquad \square$$

8.6 Let $(X, \|\cdot\|)$ be a finite-dimensional NLS. Show that there is a positive K constant such that

$$K\|\cdot\|_* \leq \|\cdot\|.$$

Solution. Let $(m_1, \ldots m_n) \in S(0,1)$ be a minimum point of f proven to exist in Problem 8.5. Let

$$K := \|\textstyle\sum_{i=1}^n m_i b_i\| = f(m_1, \ldots, m_n).$$

Then, $\quad K \leq f(\alpha_1, \ldots, \alpha_n).$

Furthermore, if $x = \sum_{i=1}^n \alpha_i b_i \in S(0,1)$, then $\|x\|_* = 1$ and

$$\|x\| = f(\alpha_1, \ldots, \alpha_n) \geq K.$$

Therefore, for each $y \in X\backslash\{\theta\}$ $\left(\text{because } \left\|\frac{y}{\|y\|_*}\right\|_* = 1\right)$,

$$\left\|\frac{y}{\|y\|_*}\right\| \geq K$$

or

$$\|y\| \geq \|y\|_* K,$$

which also holds for $y = \theta$, thus on entire X.

Because $\|\cdot\|$ and $\|\cdot\|_*$ are equivalent, any two norms on X are equivalent by transitivity (Proposition 8.5). □

8.8 Let Y be a finite-dimensional subspace of an NVS $(X, \|\cdot\|)$. Suppose there is a bounded sequence $\{y_n\} \subseteq Y$. Show that this sequence has a subsequence convergent in Y. (This generalizes a known result from calculus that any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.)

Solution. If $\{y_n\}$ is bounded, then there is a closed ball $C(\theta, M)$ such that $\{y_n\} \subseteq C(\theta, M)$. Because by Corollary 8.2, Y is closed in X and any closed ball is also closed, $Y \cap C(\theta, M)$ is closed.

Obviously, $\{y_n\} \subseteq Y \cap C(\theta, M)$. Because Y has the Heine - Borel property and $Y \cap C(\theta, M)$ is a closed and bounded subset of Y, $Y \cap C(\theta, M)$ is compact. Thus, by Theorem 6.3, $\{y_n\}$ must have a convergence subsequence in Y. □

Chapter 3. Elements of Point Set Topology

1.2 Let X be a nonempty set and $\tau = \{X, \emptyset, C^c : C \subseteq X \text{ and } C \text{ is finite}\}$. Show that τ is a topology on X. τ is called the *cofinite* (or *finite complement*) *topology* on X.

Solution. We exclude the trivial case when X is a finite set, because τ becomes a discrete topology. Then, in the general case, τ is filled with infinite sets whose complements are finite sets. Obviously, τ contains all unions of such sets. This is an easy-to-verify claim with De Morgan's law. Indeed, if $\{O_i : i \in I\} \subseteq \tau$,

$$\left(\bigcup_{i\in I} O_i\right)^c = \bigcap_{i\in I} O_i^c$$

must be a finite set as the intersection of finite sets. Furthermore, τ must be closed relative to finite intersections, because again by De Morgan,

$$\left(\bigcap_{i=1}^{n} O_i\right)^c = \bigcup_{i=1}^{n} O_i^c$$

must be finite as a union of finite sets. Then, its complement is an infinite set. □

1.4 Prove Proposition 1.6. *$A \subseteq X$ is a neighborhood for all of its points if and only if A is open.*

Solution. (i) If A is open then the statement follows immediately.

(ii) Suppose A is a neighborhood for all of its points. Then, for each $x \in A$, there is an open set $O_x \subseteq A$. Thus, $\bigcup_{x\in A} O_x$ is an open set and it is also a subset of A. On the other hand, $A = \bigcup_{x\in A} \{x\}$ and it is a subset of $\bigcup_{x\in A} O_x$, because $\{x\} \subseteq O_x$ for each x. □

1.6 Prove Proposition 1.9. *The closure of a set A is the smallest closed set containing A. Furthermore, A is closed if and only if $A = \overline{A}$.*

Solution. (i) Show that $\overline{A}$ is closed or that $\overline{A}^c$ is open. Denote $\text{Ext}A = \mathit{Int}(A^c)$. Then, $x \in \text{Ext}A \Leftrightarrow \exists U_x \subseteq A^c$. But then this U_x does not meet A and thus $U_x \subseteq \overline{A}^c$. This happens if and only if $x \in \overline{A}^c$. Because Ext$A$ is by definition an open set, then so is $\overline{A}^c$ and thus $\overline{A}$ is closed.

(ii) We show that A is closed if and only if $A = \overline{A}$. If A is closed, then obviously A is the smallest closed superset of itself and thus equals $\overline{A}$. Conversely, if $A = \overline{A}$, then A is closed.

(iii) We show that $\overline{A}$ is the smallest superset of A. Suppose there is a closed superset F of A strictly smaller than $\overline{A}$. Assume that A is not closed [or else by (ii) $\overline{A}$ is the smallest superset of A], then $A \subset F$. Then, $\mathit{Int}(A^c) = \overline{A}^c \subset F^c$. Because $\mathit{Int}(A^c)$ is the largest open subset of A^c and because $A \subset F \Rightarrow F^c \subset A^c$ we have

$$\mathit{Int}(A^c) = \overline{A}^c \subset F^c \subset A^c,$$

which shows that F^c is strictly larger than $\mathit{Int}(A^c)$, which is impossible. □

1.8 Show that $\overline{A} = A \cup \partial A$.

Solution.

1) $A^c \subseteq \overline{A^c} \Rightarrow X = A \cup A^c \subseteq A \cup \overline{A^c} = X.$

2) From 1) $\Rightarrow \overline{A} \subseteq A \cup \overline{A^c} = X \Rightarrow \overline{A} \cap [A \cup \overline{A^c}] = \overline{A}$.

3) $\overline{A} = \overline{A} \cap X$ (from 2)) $= \overline{A} \cap (A \cup \overline{A^c})$

(by the distributive law)

$= A \cup (\overline{A} \cap \overline{A^c}) = A \cup \partial A$ (by definition). □

1.10 Show that $\overline{A} = \mathring{A} + \partial A$. (*Hint*: Proceed in the same way as in Problem 2.11, Chapter 2, and work with a neighborhood instead of a ball.)

Solution.

a) We prove that $\mathring{A} \cap \overline{A^c} = \emptyset$. Let $x \in \mathring{A}$. Then, there is an open neighborhood $O_x \subseteq A \Rightarrow O_x \cap A^c = \emptyset \Rightarrow x \notin \overline{A^c}$.

b) We show that $\mathring{A} + \overline{A^c} = X$ (the carrier), which would be equivalent to $\mathring{A} = \left(\overline{A^c}\right)^c$. Obviously, $x \notin \overline{A^c} \Leftrightarrow \exists O_x \cap A^c = \emptyset \Leftrightarrow O_x \subseteq A \Leftrightarrow x \in \mathring{A}$.

c) Prove the statement by the following steps.

$$\overline{A} = \overline{A} \cap X = \overline{A} \cap (\mathring{A} + \overline{A^c})$$

$$= \mathring{A} + (\overline{A} \cap \overline{A^c}) = \mathring{A} + \partial A. \quad \square$$

1.12 Show that $\overline{A} = A \cup A'$.

Solution. Because A' and A are subsets of $\overline{A} \Rightarrow A' \cup A \subseteq \overline{A}$. On the other hand, for each $x \in \overline{A}$, each neighborhood U_x meets A, then either $x \in A$, or else $x \notin A$ but $x \in A' \Rightarrow \overline{A} \subseteq A \cup A'$. □

1.14 Show that $\partial A = \emptyset$ if and only if A is open and closed.

Solution.

1) A is open and closed implies that $A = \overline{A}$ and $A^c = \overline{A^c}$. Because $\partial A = \overline{A} \cap \overline{A^c}$, we have $\partial A = A \cap A^c = \emptyset$.

2) Let $\partial A = \overline{A} \cap \overline{A^c} = \emptyset$. From

$$\emptyset = A \cap A^c = \overline{A \cap A^c} \subseteq \overline{A} \cap \overline{A^c} = \emptyset,$$

it follows that $A \cap A^c = \overline{A} \cap \overline{A^c}$.

Now, $\overline{A} \cap A^c \subseteq \overline{A} \cap \overline{A^c} = \emptyset$. Thus, $\overline{A} \backslash A = \overline{A} \cap A^c = \emptyset$, that is, $A = \overline{A}$, for $A \subseteq \overline{A}$. Analogously, we show that $A^c = \overline{A^c}$. □

1.16 Show that the inverse inclusion in the previous problem holds if and only if A is closed and open.

Solution. Clearly,

$$\overline{A^c} \subseteq (\overline{A})^c \;\Leftrightarrow\; \overline{A^c} \backslash (\overline{A})^c = \emptyset.$$

Now,

$$\overline{A^c} \backslash (\overline{A})^c = \overline{A^c} \cap \overline{A} = \partial A = \emptyset$$

(by Problem 1.14) $\Leftrightarrow$ A is open and closed. □

1.18 A subset of a topological space (X, τ) is a G_δ set if it is a countable intersection of open sets. A subset F_σ is defined as a countable union of closed sets. Show that G_δ can be represented as a countable intersection of a monotone nonincreasing sequence of open sets. Likewise, show that F_σ can be represented as a countable union of a monotone nondecreasing sequence of closed sets.

Solution. If $F_\sigma = \bigcup_{k=1}^{\infty} F_k$, then, with $\mathcal{F}_n = \bigcup_{k=1}^{n} F_k$ we have $F_\sigma = \bigcup_{n=1}^{\infty} \mathcal{F}_n$. □

1.20 Let $\mathcal{U}_x$ be a neighborhood system at x. Show that for each $U \in \mathcal{U}_x$, there is a $V \in \mathcal{U}_x$ such that $U \in \mathcal{U}_y$ for all $y \in V$.

Solution. Given $U \in \mathcal{U}_x$, let $V = \mathring{U}$, so for each $y \in V = \mathring{U}$, U is a neighborhood of y and thus $U \in \mathcal{U}_y$. □

1.21 Let $\mathcal{B}_x$ be a neighborhood base at x. Show that if $A, B \in \mathcal{B}_x$, then there is a base neighborhood $C \in \mathcal{B}_x$ such that $C \subseteq A \cap B$.

Solution. Because $\mathcal{B}_x \subseteq \mathcal{U}_x$ and $\mathcal{U}_x$ is $\cap$-stable, by Problem 1.19, it follows that $A \cap B \in \mathcal{U}_x$. By the definition of $\mathcal{B}_x$, then $A \cap B$ is a superset to at least one base neighborhood of x that we can denote by C. □

1.25 Let $\mathcal{S} \subseteq \mathcal{P}(X)$ be a collection of subsets of a nonempty set X. Show that there is a topology containing $\mathcal{S}$. Also show that there is the weakest topology containing $\mathcal{S}$, in notation $\tau(\mathcal{S})$.

Solution. Clearly, there exists a topology containing $\mathcal{S}$ [for instance, $\mathcal{P}(X)$]. Then define $\tau(\mathcal{S})$ as the intersection of all topologies containing $\mathcal{S}$. We show that $\tau(\mathcal{S})$ is a topology on X.

(i) X and Ø belong to all topologies containing $\mathcal{S}$. Therefore, X and Ø belong to $\tau(\mathcal{S})$.

(ii) Let $O_1, O_2, \ldots, O_n \in \tau(\mathcal{S})$. Then $O_1, O_2, \ldots, O_n$ are elements of every topology containing $\mathcal{S}$. This implies that $\bigcap_{k=1}^{n} O_k$ belongs to all topologies containing $\mathcal{S}$, and thus it belongs to $\tau(\mathcal{S})$.

(iii) By similar arguments, $\tau(\mathcal{S})$ is closed relative to the formation of arbitrary unions.

Obviously, $\tau(\mathcal{S})$ is the weakest topology containing $\mathcal{S}$. □

2.2 Show that $\mathcal{B} \subseteq \mathcal{P}(X)$ is a base for a topology on X if and only if two conditions are met:

(i) Each $x \in X$ belongs to at least one set $B \in \mathcal{B}$ (or equivalently, $X = \bigcup_{B \in \mathcal{B}} B$)

(ii) $\forall B_1, B_2 \in \mathcal{B}$ and $\forall x \in B_1 \cap B_2$, $\exists B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \cap B_2$.

[*Hint*: Use the steps that follow.

1) If $\mathcal{B}$ is a base, then apply Theorem 2.2 (ii).

2) Let $\tau = \left\{ \bigcup_{B \in \mathcal{B}'} B : \forall\, \mathcal{B}' \subseteq \mathcal{B} \right\}$. Show that τ is a topology on X and that $\mathcal{B}$ is a base for τ.]

Solution.

1) Let $\mathcal{B}$ be a base for (X, τ). Then statement (i) follows from the fact that each open set, in particular, X, can be represented as the union of base sets. Statement (ii) is due to the following. Because $B_1, B_2 \in \mathcal{B}$, $B_1 \cap B_2 = \bigcup_i B_i'$, with $B_i' \in \mathcal{B}$. Thus, for an $x \in B_1 \cap B_2$, there is a $B \in \{B_i'\}$ such that $B \subseteq B_1 \cap B_2$.

2) Let (i) and (ii) hold and let

$$\tau = \left\{ \bigcup_{B \in \mathcal{B}'} B : \forall\, \mathcal{B}' \subseteq \mathcal{B} \right\}.$$

Then, τ is a topology on X and $\mathcal{B}$ is a base for τ. Indeed, from (i) it follows that $X \in \tau$, also τ is closed relative to the formation of all unions. It remains to show that the intersection of any two elements of τ belongs to τ. If $O_1, O_2 \in \tau$, then $O_k = \cup\, B^k$, $k = 1, 2$, for some $\{B^k\} \subseteq \mathcal{B}$ (the index k is not the one over which the union is formed). Let

$$x \in O_1 \cap O_2 = (\cup\, B^1) \cap (\cup\, B^2).$$

Therefore, there is at least one set $B' \in \{B^1\}$ and one set $B'' \in \{B^2\}$ such that $x \in B' \cap B''$. By assumption (ii),

$$\exists\, B_x \subseteq B' \cap B'' \colon x \in B_x \in B' \cap B'' \text{ and } B_x \in \mathcal{B}.$$

Hence for each $x \in O_1 \cap O_2$, there is a $B_x \in \mathcal{B}$ such that $B_x \subseteq O_1 \cap O_2$, which implies that the latter is an open set, and

$$O_1 \cap O_2 = \bigcup_{x \in O_1 \cap O_2} B_x.$$

Thus, τ is a topology on X. The fact that $\mathcal{B}$ is a base for τ follows directly from the construction of τ. □

2.4 Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be bases for the topologies (X, τ_1) and (X, τ_2), respectively. Prove the equivalence of the following statments:

- (i) For each $x \in X$ and each base set $A_x \in \mathcal{B}_1$ that contains x, there is a base set $B_x \in \mathcal{B}_2$ which contains x and $B_x \subseteq A_x$.
- (ii) τ_2 is finer than τ_1.

Solution. $(i) \Rightarrow (ii)$. Let $B_1 \in \mathcal{B}_1$ and let $x \in B_1$ be any element of B_1. According to assumption (i), there is a base set $B_{2x} \in \mathcal{B}_2$ that also contains x and is a subset of B_1. Thus, any B_1 can be "filled" with base elements from $\mathcal{B}_2$, that is $B_1 = \bigcup_{x \in B_1} B_{2x}$ which makes B_1 a τ_2 element. In other words, $\mathcal{B}_1 \subseteq \tau_2$. By Theorem 2.4 (i), then $\tau_1 \subseteq \tau_2$.

$(ii) \Rightarrow (i)$. Let $x \in X$. Then, there is a base element $B_1 \in \mathcal{B}_1$ that contains x. Because $\tau_1 \subseteq \tau_2 \Rightarrow B_1 \in \tau_2$. Now being an open set in τ_2, B_1 can be represented by a union of base sets from $\mathcal{B}_2$, at least one of which contains x. Thus, there is a $B_2 \subseteq B_1$ that contains x. □

2.6 Show that the collection of all sets on the real line of types (a, ∞) and $(-\infty, b)$ is a subbase for the usual topology $(\mathbb{R}, \tau_e)$.

Solution. Obviously the union of all intervals of type (a, ∞) and $(-\infty, b)$ gives $\mathbb{R}$. Thus, their collection $\mathcal{S}$ is a subbase for $\tau(\mathcal{S})$. Now, if we extend $\mathcal{S}$ by including in $\mathcal{S}$ all finite intersections, we arrive at the base $\mathcal{B}$ for $\tau(\mathcal{S})$, in accordance with Theorem 2.8. On the other hand, $\mathcal{B}$ will include intervals (a, b) which are known to be base sets for τ_e (the usual topology). Therefore, τ_e and $\tau(\mathcal{S})$ share the same base sets and hence must coincide. The latter yields that $\mathcal{S}$ is a subbase for τ_e. □

2.8 Show that the usual topology τ_e on $\mathbb{R}^n$ and the product topology τ_P generated by $\underset{i=1}{\overset{n}{\times}} \tau_e^{(i)}$ coincide.

Solution. By Example 2.3(iii), $\mathcal{R}$ (the set of all base rectangles in $\mathbb{R}^n$) is a base for $(\mathbb{R}^n, \tau_e)$. Recall that a rectangle of type $R = \prod_{i=1}^{n} (a_i, b_i)$ is a *base* or *simple* rectangle. In other words, $\tau(\mathcal{R}) = \tau_e$.

On the other hand,

$$\mathcal{R} \subseteq \tau_e \times \cdots \times \tau_e \subseteq \tau_P = \tau(\tau_e \times \cdots \times \tau_e).$$

Thus, by Theorem 2.4 (i), $\tau(\mathcal{R}) = \tau_e \subseteq \tau_P$. In addition, the generator of τ_P,

$$\tau_e \times \cdots \times \tau_e \subseteq \tau_e.$$

Indeed, O_i is a union of open intervals and equals $\bigcup(a_i, b_i)$. Thus, by Problem 5.2(ii), Chapter 1,

$$O_1 \times O_2 = \bigcup(O_1 \times I_2)$$

$$= \bigcup\left(\left[\bigcup I_1\right] \times I_2\right) = \bigcup\bigcup(I_1 \times I_2).$$

By induction, $O_1 \times \cdots \times O_n$ is a union of open base rectangles and thus is open in τ_e.

Having shown this, if we use

$$\tau_1 = \tau_e, \tau_2 = \tau_P,$$

$$\mathcal{B}_1 = \mathcal{R}, \text{ and } \mathcal{B}_2 = \tau_e \times \cdots \times \tau_e,$$

we have by Theorem 2.4(iii) that

$$\tau_1 = \tau_e = \tau_2 = \tau_P. \qquad \square$$

3.1 Prove Proposition 3.5. *Let* (X, τ) *be a Hausdorff topological space,* $\lim_{n\to\infty} x_n = x$, *and let* $\lim_{n\to\infty} x_n = y$. *Then* $x = y$.

Solution. $x_n \to x \Leftrightarrow \forall$ neighborhood U_x, $\exists k$ such that

$$\{x_k, x_{k+1}, \dots\} \subseteq U_x.$$

$x_n \to y \Leftrightarrow \forall$ neighborhood U_y, $\exists m$ such that

$$\{x_m, x_{m+1}, \dots\} \subseteq U_y.$$

If $x \neq y$ then $\exists$ U_x and U_y such that $U_x \cap U_y = \emptyset$. Let $s = \max\{k, m\}$. Then,

$$\{x_s, x_{s+1}, \dots\} \subseteq U_x \text{ and } \{x_s, x_{s+1}, \dots\} \subseteq U_y,$$

which is impossible because $U_x \cap U_y = \emptyset$. $\square$

3.3 Show that any metric space is first countable.

Solution. By Example 2.3(ii), $\mathcal{B}_x = \{B(x, q)\colon\ q \in \mathbb{Q},\ q > 0\}$ is a neighborhood base at $x \in X$ and it is countable. $\square$

3.5 Is it true that any first countable topological space is also second countable?

Solution. No, it is not. For example, $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ is the discrete topological space. This is first countable (specifically, because it is metrizable). However, it is not second countable. For each x, $\{x\}$

must be a base set, and obviously, $\mathcal{B} = \{\{x\}\colon x \in \mathbb{R}\}$ would be the "smallest" base for $\mathcal{P}(\mathbb{R})$ and it is not countable. □

3.7 Show that subspaces of second countable spaces are second countable.

Solution. Let (X,τ) be second countable, let (Y,τ') be a subspace of (X,τ) and let $\mathcal{B}$ be a countable base for τ. We will show that $\mathcal{B}' := \mathcal{B} \cap Y := \{B \cap Y\colon B \in \mathcal{B}\}$ is a base for (Y,τ') and that the latter is second countable. If $O' \in \tau'$, then $O' = O \cap Y$ for some $O \in \tau$. Now, because

$$O = \bigcup_{\text{some } B_i \in \mathcal{B}} B_i,$$

we have that

$$O' = \bigcup_{\text{some } B_i \cap Y \in \mathcal{B}'} (B_i \cap Y). \qquad \square$$

4.3 Prove Proposition 4.4. *Let (X,τ), (Y,τ_1), and (Z,τ_2) be topological spaces and let $f\colon X \to Y$ and $g\colon Y \to Z$ be continuous functions. Then the function $g \circ f\colon X \to Z$ is continuous.*

Solution. Let $H \in \tau_2$. Then by Theorem 4.2, $g^*(H) \in \tau_1$, that is the inverse image of H is open in Y. Again, applying Theorem 4.2, we have $f^* \circ g^*(H) \in \tau$, now open in X. Finally, recalling that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ we complete the proof of the proposition. □

4.5 Under the conditions of Problem 4.4, set $f(x) = x + 1$. Is f continuous?

Solution. It is easily seen that $f^*(\mathbb{R}) = \mathbb{R}$, $f^*(\emptyset) = \emptyset$, $f^*(y) = y - 1 \Rightarrow f^*(\{2\}) = \{1\} \in \tau$, $f^*([2,4)) = [1,3) \in \tau \Rightarrow f$ is continuous. □

4.7 Under the condition of Problem 4.6, assume that $\tau' = \tau(d)$, that is, (Y,τ') is a metrizable topological space.

a) Show that f is continuous at $x \in X$ if and only if the inverse image under f of any open ball $B_d(f(x), \varepsilon)$ is a neighborhood of x.

b) Show that, for each open ball $B_d(f(x), \varepsilon)$ there is a neighborhood $U_x(\varepsilon)$ such that

$$f(U_x(\varepsilon)) \subseteq B_d(f(x), \varepsilon)\,.$$

Solution. a) The statement of this problem follows from Problem 4.6 as a special case where any ball centered at $f(x)$ belongs to a neighborhood base at $f(x)$.

b) By a), $f^*(B_d(f(x), \varepsilon))$ is a neighborhood of x. By Problem 2.6 (Chapter 1),

$$f_* \circ f^*(B_d(f(x), \varepsilon)) \subseteq B_d(f(x), \varepsilon).$$

Thus, $f^*(B_d(f(x), \varepsilon)$ is such an appropriate neighborhood $U_x(\varepsilon)$ and b) follows. □

4.8 Let $f\colon (X, \tau) \to (Y, \|\cdot\|)$ be a map, where Y is an NVS over a field $\mathbb{F}$. Show that f is continuous at $x \in X$ if and only if, for every $\varepsilon > 0$, there is a neighborhood $U_x(\varepsilon) \in \mathcal{U}_x$ such that for each $y \in U_x(\varepsilon)$, $\|f(x) - f(y)\| < \varepsilon$.

Solution. This is a corollary to statement in Problem 4.7 b). Indeed, 4.7 b) is equivalent to the statement: "$\forall\ y \in U_x(\varepsilon)$, it holds $d(f(x), f(y)) < \varepsilon$." Because d is TIH,

$$d(f(x), f(y)) = d(f(x) - f(y), \theta)$$

$$= \| f(x) - f(y) \|_d < \varepsilon.$$ □

4.11 Prove Proposition 4.14. *Let $f, g\colon (X, \tau) \to (Y, \tau')$ be two continuous maps that coincide on some dense set in X. If (X, τ) is first countable and if (Y, τ') is Hausdorff, then $f = g$ on X.*

Solution. By Lemma 4.13, the set $S = \{x\colon f = g\}$ is closed, because S is dense in $X \Rightarrow S = \overline{S} = X$. □

4.13 Show that $(\mathbb{R}, \tau_e)$ is homeomorphic to $(-1, 1)$ with the corresponding relative topology on $(-1, 1)$.

Solution. The function $f(x) = \tan(\frac{2}{\pi}x)$ is a relevant homeomorphism between $(\mathbb{R},\tau_e)$ and $(-1,1)$. □

5.1 Let (Y_1, τ_1) and (Y_2, τ_2) be topological spaces, let (Ω, τ_P) be the product space, and let $F_i \in \tau_i^c, i = 1, 2$. Show that $F_1 \times F_2 \in \tau_P^c$.

Solution. By Problem 5.1(ii) of Chapter 1, $(F_1 \times F_2)^c = (F_1^c \times Y_2) \cup (Y_1 \times F_2^c)$ implying that the latter is the union of two open base sets in τ_P. Thus, $F_1 \times F_2$ is closed in τ_P. □

5.3 Show that $\mathit{Int}(A \times B) = \mathring{A} \times \mathring{B}$.

Solution. (i) Clearly, $\mathring{A} \times \mathring{B}$ is open in τ_P and because $\mathring{A} \subseteq A$ and $\mathring{B} \subseteq B$, it follows that $\mathring{A} \times \mathring{B} \subseteq A \times B$. But $\mathit{Int}(A \times B)$ is the largest open subset of $A \times B$, whereas $\mathring{A} \times \mathring{B}$ is just an open subset of $A \times B$. Thus, $\mathring{A} \times \mathring{B} \subseteq \mathit{Int}(A \times B)$.

(ii) The inverse inclusion. Because $\mathit{Int}(A \times B)$ is open in τ_P, for any point $\xi = (x, y) \in \mathit{Int}(A \times B)$, there is an open neighborhood O_ξ which is in $\mathit{Int}(A \times B)$. As an open set in τ_P,

$$O_\xi = \bigcup_{O_1 \in \tau_1, O_2 \in \tau_2} O_1 \times O_2.$$

Therefore, ξ belongs to one of these open sets, say $O_1 \times O_2$. Hence, $x \in O_1$ and $y \in O_2$. Obviously, $O_1 \subseteq A$ and $O_2 \subseteq B$. Thus, x and y are interior points of A and B, that is, $\xi \in \mathring{A} \times \mathring{B}$. □

5.6 Let $(Y_i, \tau_i), i = 1, \ldots, n,$ be topological spaces and (Ω, τ_P) be the product space. If $A_i \subseteq Y_i, i = 1, \ldots, n$, show that

$$\overline{\prod_{i=1}^{n} A_i} = \prod_{i=1}^{n} \overline{A}_i.$$

Solution. (i) We first show that

$$\overline{\prod_{i=1}^{n} A_i} \subseteq \prod_{i=1}^{n} \overline{A}_i.$$

One can show part (i) in exactly the same way as in part (i) of Problem 5.21. Alternatively, one can show it through the following steps.

Step 1. If f is a continuous function, then $\overline{f^*(B)} \subseteq f^*(\overline{B})$. Indeed, $f^*(B) \subseteq f^*(\overline{B})$ and $f^*(\overline{B})$ is a closed superset of $f^*(B)$, whereas $\overline{f^*(B)}$ is the smallest closed superset of $f^*(B)$.

Step 2. Recall from Problem 1.7 that $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$, which can be extended by induction for any finitely many sets.

Step 3. From Step 2 and then Step 1, Because the projection maps π_i's are continuous, it follows that

$$\overline{\bigcap_{i=1}^{n} \pi_i^*(A_i)} \subseteq \bigcap_{i=1}^{n} \overline{\pi_i^*(A_i)} \text{ (Step 1) } \subseteq \bigcap_{i=1}^{n} \pi_i^*(\overline{A}_i),$$

thereby concluding part (i).

(ii) Show the inverse inclusion. Let

$$\xi = (y_1, \ldots, y_n) \in \prod_{i=1}^{n} \overline{A}_i.$$

Thus, $y_i \in \overline{A}_i$. Let O_ξ be any open neighborhood of ξ in Ω. Therefore, it can be formed by a union of open base rectangles of type $\prod_{i=1}^{n} O_i$. Because ξ belongs to the union, it must belong to one of these

sets, say to $\prod_{i=1}^{n} O_{y_i}$. Consequently, O_{y_i} is an open neighborhood of y_i in τ_i and because $y_i \in \overline{A}_i$, it follows that O_{y_i} must meet A_i. Thus, $\prod_{i=1}^{n} O_{y_i}$ must meet $\prod_{i=1}^{n} A_i$ implying that ξ is a closure point of $\prod_{i=1}^{n} A_i$. The latter concludes part (ii). □

5.8 Prove Proposition 5.1. *Let*

$$\widetilde{\mathcal{B}} = \mathcal{B}_1 \times \cdots \times \mathcal{B}_n : = \{B_1 \times \cdots \times B_n : B_i \in \mathcal{B}_i\}$$

where $\mathcal{B}_i$ is a base for τ_i, $i = 1, \ldots, n$. Then $\widetilde{\mathcal{B}}$ is also a base for τ_P.

Solution. Let $O \in \tau_P$. Because $\mathcal{B} = \{O_1 \times \cdots \times O_n : O_i \in \tau_i, i = 1, \ldots, n\}$ is a base for τ_P,

$$O = \bigcup_{j \in J} \mathcal{O}^{[j]},$$

where $\mathcal{O}^{[j]} = O_1^{[j]} \times \cdots \times O_n^{[j]} \in \mathcal{B}$ and $O_s^{[j]} \in \tau_s$, $s = 1, \ldots, n$.

If $\mathcal{B}_s$ is a base for τ_s, then we have

$$O_s^{[j]} = \bigcup_{i_s \in I_s} B_{i_s}^{[j]} \text{ with } B_{i_s}^{[j]} \in \mathcal{B}_s,\ s = 1, \ldots, n.$$

Now, taking into account that

$$O_1^{[j]} \times \cdots \times O_n^{[j]} = \bigcap_{s=1}^{n} \pi_s^* \left(O_s^{[j]} \right)$$

by the distributive law, we have that

$$O = \bigcup_{j \in J} \mathcal{O}^{[j]} = \bigcup_{j \in J} \left[\bigcap_{s=1}^{n} \pi_s^* \left(\bigcup_{i_s \in I_s} B_{i_s}^{[j]} \right) \right]$$

$$= \bigcup_{j\in J} \bigcup_{i_1\in I_1} \dots \bigcup_{i_n\in I_n} \left\{ \pi_1^*\left(B_{i_1}^{[j]}\right) \cap \dots \cap \pi_n^*\left(B_{i_n}^{[j]}\right) \right\}$$

$$= \bigcup_{j\in J} \bigcup_{i_1\in I_1} \dots \bigcup_{i_n\in I_n} \left(B_{i_1}^{[j]} \times \dots \times B_{i_n}^{[j]} \right).$$

The latter shows that every open set from τ_P can be represented as the union of base parallelepipeds. □

5.10 Prove Proposition 5.3. *Let*

$$\mathcal{S}' = \{\pi_i^*(S_i) : S_i \in \mathcal{S}_i,\ i = 1, \dots, n\},$$

where $\mathcal{S}_i$ is a subbase for τ_i. Then $\mathcal{S}'$ is a subbase for $\tau_P = \tau(\mathcal{S})$.

Solution. In equation (P5.2) we replace $\prod_{k=1}^{n} S_{i_k}^{(k)}$ with $\bigcap_{k=1}^{n} \pi_k^*\left(S_{i_k}^{(k)}\right)$. Then $\widetilde{\mathcal{B}}$ is constructed as in Problem 5.1 (from those $\mathcal{B}_i$'s) is a base for τ_P. □

5.12 Let $f : (\Omega, \tau) \to \left(\Omega = \prod_{i=1}^{n} Y_i, \tau_P\right)$ and $\pi_j : \Omega \to Y_j$. Show that the function f is continuous if and only if each $\pi_j \circ f$ is continuous. [*Hint*: Show that $f^*(S) \in \tau$, for every subbase element of τ_P, and then apply Theorem 4.8.]

Solution.

1) If f is continuous then $\pi_j \circ f$ is also continuous, for the composition preserves continuity.

2) Let each $\pi_j \circ f$ be continuous and let $S = S_1 \times \dots \times S_n$ be a subbase set in τ_P. Then

$$f^*(S) = f^*\left(\bigcap_{k=1}^{n} \pi_k^*(S_k)\right) = \bigcap_{k=1}^{n} (\pi_k \circ f)^*(S_k).$$

Because $\pi_k \circ f$ is continuous, it follows that $(\pi_k \circ f)^*(S_k) \in \tau$, $k = 1, \ldots, n \;\Rightarrow\; f^*(S) \in \tau$. By Theorem 4.8 it is equivalent to continuity of f, for we showed that $f^*(S) \in \tau$ for an arbitrary subbase set. □

5.14 Let $(Y_x, \tau_x), x \in X,$ be an indexed family of topological spaces and (Ω, τ_P) be the Tychonov product space. Let A be an arbitrary subset of X and $F_x \in \tau_x^c, x \in A$. Show that $\prod\limits_{x \in A} F_x$ is closed in τ_P.

Solution.

$$\left(\prod_{x \in A} A_x\right)^c = \left(\bigcap_{x \in A} \pi_x^*(F_x)\right)^c = \bigcup_{x \in A} \pi_x^*(F_x^c),$$

which is a union of open sets. In fact, each $\pi_x^*(F_x^c)$ is a base open set in τ_P. □

5.16 Let $(Y_x, \tau_x), x \in X,$ be an indexed family of topological spaces such that $\mathcal{S}_x$ is a subbase for $\tau_x, x \in X$. Show that the collection of all unit subbase cylinders

$$\mathcal{S} = \bigcup_{x \in X} \pi_x^{**}(\mathcal{S}_x) = \{\pi_x^*(S_x) : S_x \in \mathcal{S}_x, x \in X\}$$

[in Remark 5.8(i) and Remark 5.10] is a subbase for the Tychonov topology.

Solution. 1) Suppose $\mathcal{S}_x$ is a subbase for the factor space (Y_x, τ_x) and $\mathcal{B}_x$ is the associated base for τ_x in Y_x. According to Problem 5.15, simple cylinders such as $\pi_{x_1}^*\left(B_{x_1}\right) \cap \cdots \cap \pi_{x_n}^*\left(B_{x_n}\right)$ form a base $\widetilde{\mathcal{B}}$ for Tychonov topology τ_P. To simplify the notation let $B \in \widetilde{\mathcal{B}}$ be in the form

$$B = \pi_1^*(B_1) \cap \cdots \cap \pi_n^*(B_n). \tag{P5.16}$$

Now, because B_y is a base set in τ_y, it can be represented through subbase elements as

$$B_y = \bigcap_{i_y \in I_y} S_{i_y}^{(y)},\ y = 1, \ldots, n,$$

where I_y is a finite subset of $\mathbb{N}$, i_y means a variable index associated with $y,$ and $S_{i_y}^{(y)} \in \mathcal{S}_y$. Thus, from (P5.16) we have

$$B = \pi_1^* \left(\bigcap_{i_1 \in I_1} S_{i_1}^{(1)} \right) \cap \cdots \cap \pi_n^* \left(\bigcap_{i_n \in I_n} S_{i_n}^{(n)} \right)$$

$$= \bigcap_{i_1 \in I_1} \cdots \bigcap_{i_n \in I_n} \left\{ \pi_1^* \left(S_{i_1}^{(1)} \right) \cap \cdots \cap \pi_n^* \left(S_{i_n}^{(n)} \right) \right\}. \tag{P5.16a}$$

So, B ends up being a finite intersection of simple cylinders of the form

$$\pi_1^* \left(S_{i_1}^{(1)} \right) \cap \cdots \cap \pi_n^* \left(S_{i_n}^{(n)} \right) \tag{5.16b}$$

that are made of subbase sets. Thus, $B = \prod_{x \in X} R_x,$ where

$$R_x = \begin{cases} S_x, & x \in A \subseteq X \\ Y_x, & x \in A^c \end{cases}$$

with A as a finite subset of X and S_x as a subbase element of $\mathcal{S}_x \subseteq \tau_x$. According to Theorem 2.8, finite intersections of subbase elements (simple cylinders) of type (P5.16b), form a base $\mathcal{B}$ for the weakest topology $\tau(\mathcal{S})$. Because $\mathcal{B} = \mathcal{B}(\mathcal{S})$ and because τ_P and $\tau(\mathcal{S})$ share the same base, by Theorem 2.4(ii), they coincide.

Notice that simple cylinders are made of unit cylinders. The latter can be utilized as more rudimentary subbase elements. Finite intersections of them as in (P5.16b) lead us to base $\mathcal{B}$ through

(P5.16a). Thus unit cylinders also serve as a subbase for τ_P. In other words,

$$\mathcal{S} = \bigcup_{x\in X} \pi_x^{**}(\mathcal{S}_x)$$

is a subbase for τ_P.

Finally, we have

$$\Omega = \pi_x^*(Y_x) = \pi_x^*\left(\bigcup_{i\in I} S_i^{(x)}\right) = \bigcup_{i\in I} \pi_x^*\left(S_i^{(x)}\right)$$

proving that $\mathcal{S}$ indeed is a subbase for τ_P. □

5.18 Let $\left(\Omega = \prod_{x\in X} Y_x, \tau_P\right)$ be the Tychonov topology and assume that each factor space is first countable. Is (Ω, τ_P) first countable if:

a) $|X| = \aleph_0$?
b) $|X| \geq \mathfrak{c}$?

Solution. Let $O_\xi \in \tau_P$ be an open neighborhood of a point $\xi \in \Omega$. Because

$$\mathcal{B} = \left\{\pi_A^*\left(\prod_{x\in A} O_x\right) : X \supseteq A \prec \mathbb{N}\right\}$$

is a base for τ_P, O_ξ is representable in the form $\bigcup_{B\in\mathcal{B}'} B$ for $\mathcal{B}' \subseteq \mathcal{B}$. Thus, $\xi \in$ some B_ξ of this union which is a simple cylinder of the form

$$B_\xi = \pi^*_{\{x_1,\ldots,x_n\}}\left(\prod_{i=1}^n O_{x_i}\right) = \pi^*_{x_1}(O_{x_1}) \cap \cdots \cap \pi^*_{x_n}(O_{x_n}).$$

Now, because each Y_x is first countable, there is a countable neighborhood base $\mathfrak{B}_{\xi(x)} = \left\{A^{(1)}_{\xi(x)}, A^{(2)}_{\xi(x)}, \dots\right\}$ at the x-coordinate $\xi(x)$ of ξ such that there is some $A^{(N)}_{\xi(x)} \subseteq O_x$ for any neighborhood O_x of $\xi(x)$. Repeating the same process for each $x_i \in \{x_1, \dots, x_n\}$, we have that $\pi^*_{\{x_1,\dots,x_n\}}\left(\prod_{i=1}^{n} A^{(N_i)}_{\xi(x_i)}\right)$ is a neighborhood of ξ which is a subset of B_ξ and thus of O_ξ.

Now we form the neighborhood base at ξ as follows. Let $\mathfrak{B} = \left\{\mathfrak{B}_{\xi(x)} : x \in X\right\}$ and let $\mathcal{B}_\xi$ be the set of all finite Cartesian products of sets from $\mathfrak{B}$. If X is countable, then according to Problem 6.14, Chapter 1, $\mathcal{B}_\xi$ is countable. From the construction of $\mathcal{B}_\xi$ and due to the nature of the Tychonov topology it follows that $\mathcal{B}_\xi$ is a neighborhood base at ξ and it is countable if so is X. Indeed, because the choice of O_ξ was arbitrary, we could pick out a finite set $\left\{A^{(N_i)}_{\xi(x_i)}, i = 1, \dots, n\right\}$ from neighborhood bases at each pertinent coordinate $\xi(x)$, with $x \in A$.

However, $\mathcal{B}_\xi$ is not countable if X is not countable. □

5.20 Let $\left(\Omega = \prod_{x\in X} Y_x, \tau_P\right)$ be the Tychonov topology and assume that each factor space is second countable. Is (Ω, τ_P) second countable if X is countable?

Solution. Yes, Ω is second countable. According to Problem 5.15, τ_P can be generated by the base

$$\widetilde{\mathcal{B}} = \left\{\pi^*_A\left(\prod_{x\in A} B_x\right) : A\text{'s are finite subsets of } X,\ B_x \in \mathfrak{B}_x\right\}$$

where $\mathfrak{B}_x$ is a countable base for τ_x. Therefore, $\widetilde{\mathcal{B}}$ is induced by the set

$$\mathfrak{A} : = \left\{\underset{x\in A}{\times} \mathfrak{B}_x : A \subseteq X, A \prec \mathbb{N}\right\}$$

in agreement with the following specification. Given an $n \in \mathbb{N}$, a set $A = \{x_1, \ldots, x_n\} \in X$ is selected. According to A, n factor spaces $Y_{x_1}, \ldots, Y_{x_n}$ are chosen and the corresponding bases $\mathfrak{B}_{x_1}, \ldots, \mathfrak{B}_{x_n}$. From these bases we pick out respective elements $B_{x_1}, \ldots, B_{x_n}$, one from each followed by the formation of the cylinder

$$B = \pi^*_{\{x_1,\ldots,x_n\}}\left(\prod_{i=1}^{n} B_{x_i}\right) = \pi^*_{x_1}(B_{x_1}) \cap \cdots \cap \pi^*_{x_n}(B_{x_n}) \in \widetilde{\mathcal{B}}.$$

The selection process described above is identical to that in Problem 6.14, Chapter 1, according to which the set $\mathfrak{A}$ must be countable. $\square$

5.22 Let $\left(\Omega = \prod_{x\in X} Y_x, \tau_P\right)$ be the Tychonov topology and assume that each factor space is separable. Is (Ω, τ_P) separable if X is countable?

Solution. Yes, it is. Let D_x be a dense set in (Y_x, τ_x) and consider the set $\{D_x : x \in X\}$. Let

$$\{D_{x_1} \times \cdots \times D_{x_n} : n \in \mathbb{N}, \{x_1, \ldots, x_n\} \subseteq X\}$$

be the set of all finite Cartesian products from $\{D_x : x \in X\}$. Now, if $D_{x_1} \times \cdots \times D_{x_n}$ is picked out, then we select a point $\{d_{x_1}, \ldots, d_{x_n}\} \in D_{x_1} \times \cdots \times D_{x_n}$ and include it in set $\mathfrak{D}$. According to Problem 6.14, Chapter 1, the set of all such points is countable. Finally, we define the set $\mathcal{D}$ as follows. Let $\mathbf{d} \in \prod_{x\in X} D_x$ be a fixed arbitrary point from this product. Given a set $A = \{x_1, \ldots, x_n\} \subseteq X$ and the associated point $\{d_{x_1}, \ldots, d_{x_n}\} \in \mathfrak{D}$, we complement $\{d_{x_1}, \ldots, d_{x_n}\}$ to a point in Ω by the elements d_x from $\mathbf{d}$, with $x \in A^c$ and denote such a point by $\mathbf{d}_A$. Consequently,

$$\mathcal{D} = \{\mathbf{d}_A : A = \{x_1, \ldots, x_n\} \subseteq X, n \in \mathbb{N}\} \approx \mathfrak{D}$$

and thus $\mathcal{D}$ is countable.

Now we show that $\overline{\mathcal{D}} = \Omega$. To prove this statement we need to verify that for each $\xi \in \Omega$, any neighborhood $\mathcal{O}_\xi$ meets $\mathcal{D}$. As usual, there is a base neighborhood O being a subset of $\mathcal{O}_\xi$ in the form

$$O = \pi^*_{\{x_1,\ldots,x_n\}}\left(\prod_{i=1}^{n} O_{\xi(x_i)}\right), \text{ where } O_{\xi(x_i)} \in \tau_{x_i}$$

and being a neighborhood of ξ. We recall that $\xi(x)$ is the xth coordinate of ξ in the factor space Y_x. Now, because D_x is dense in Y_x, $O_{\xi(x)}$ meets D_x for each $x \in X$, in particular, for each $x \in A = \{x_1, \ldots, x_n\}$. Therefore, each $O_{\xi(x_i)}$ contains at least one point of D_{x_i}, say d_{x_i}. Therefore, $\prod_{i=1}^{n} O_{\xi(x_i)}$ meets the set $\{d_{x_1}, \ldots, d_{x_n}\}$ which by our construction must be included in $\mathfrak{D}$ and therefore, O includes the point $\mathbf{d}_A$ and thus meets $\mathcal{D}$. Consequently, $\mathcal{O}_\xi$ meets $\mathcal{D}$. □

6.1 Is the property of a set to be dense vaguely hereditary? (Consider Example 6.4.)

Solution. Although we know from Example 6.4 that the property of a set to be dense is not hereditary, we can show that it is vaguely hereditary. Let Y be open in (X,τ) and let D be a dense subset of X. Recall that then [cf. Definitions 1.8(i-ii)] for each $x \in X$ and any neighborhood U_x of x, $U_x \cap D \neq \emptyset$.) We claim that $D \cap Y$ is dense in Y. If this is not the case, then there is a $y \in Y$ and an open neighborhood U_y of y in (X, τ) such that $U_y \cap Y \cap D = \emptyset$ (as we have seen it in Example 6.4) and with $U_y \cap Y$ being a neighborhood of y in the subspace (Y, τ_Y).

On the other hand, because $U_y \cap Y$ remains an open neighborhood of y in the original space and this is due to our assumption that Y is open in (X, τ), it contradicts the very property of D being dense in (X, τ). Now it is obvious what exactly we missed in Example 6.4 regarding $U_y \cap Y$: the latter (in contrast to our present case) failed to be a neighborhood of y in the original space (X, τ). □

6.3 Show that the topological space in Problem 6.2 is not second countable.

Solution. According to Problem 6.2, $(X, \mathcal{P}(X))$ is not separable. By Problem 3.6, a nonseparable space cannot be second countable. □

6.5 Prove that the Moore plane is not second countable.

Solution. To see why the Moore plane is not second countable, we observe that according to Problem 6.3, any discrete topological space with an uncountable carrier is not second countable. Because second countability is hereditary, the Moore plane is thus not second countable. □

6.7 Let τ be the cofinite topology on an arbitrary nonempty set X. Show that (X, τ) is compact.

Solution. Recall that (X, τ) is cofinite if $\tau = \{X, \emptyset, C^c : C \prec \mathbb{N}\}$. To avoid triviality, we assume that X is an infinite set. Let $\mathcal{C}$ be an open cover of X and let some $C^c \in \mathcal{C}$. Then, because C^c also belongs to τ, by definition, C is finite, say, $\{x_1, \ldots, x_n\}$. Now, there is yet another set $C_1^c \in \mathcal{C}$, which contains at least one point, say x_1. (If no such set of $\mathcal{C}$ is available, then $\mathcal{C}$ is not a cover of X.) Then, $C \cap C_1 \subseteq \{x_2, \ldots, x_n\}$. Continuing with this process, it is easy to see that there is a $C_n^c \in \mathcal{C}$ such that $C \cap C_1 \cap \ldots \cap C_n = \emptyset$ and thus $\{C^c, C_1^c, \ldots, C_n^c\}$ is a finite subcover of X, which proves that X is compact. □

6.9 Let $\{M_1(k), \ldots, M_{j_k}(k), S_k\}$ be a cover of a set X for each $k = 1, \ldots, n$. Show that $\{M_1(k), \ldots, M_{j_k} \colon k = 1, \ldots, n\,,$ and $\bigcap_{k=1}^{n} S_k\}$ is also a cover of X.

Solution. By the assumption, $X \subseteq \left(\bigcup_{i=1}^{j_k} M_i(k) \right) \cup S_k$. Thus,

$$X \subseteq \bigcap_{k=1}^{n} \left[\left(\bigcup_{i=1}^{j_k} M_i(k) \right) \cup S_k \right]$$

$$= \left(\bigcap_{k=1}^{n} \bigcup_{i=1}^{j_k} M_i(k) \right) \cup \left(\bigcap_{k=1}^{n} S_k \right)$$

$$\subseteq \left(\bigcup_{k=1}^{n} \bigcup_{i=1}^{j_k} M_i(k) \right) \cup \left(\bigcap_{k=1}^{n} S_k \right). \qquad \square$$

6.11 Let O be an open set in $\mathbb{R}^n$. Show that there is a monotone increasing sequence $\{O_k\}\uparrow$ of open bounded subsets of O such that $\{O_k\}\uparrow O$.

Solution. Obviously, $\forall\ x \in O$, there is $B(x, \varepsilon_x) \subseteq O$, and hence $O = \bigcup_{x \in O} B(x, \varepsilon_x)$. Because $\mathbb{R}^n$ is Lindelöf and so is $(O, \tau_e \cap O)$ (vaguely hereditary), the above cover, $\{B(x, \varepsilon_x)\colon x \in O\}$, of O can be reduced to at most countable subcover, i.e.,

$$O = \bigcup_{k=1}^{\infty} B(x, \varepsilon_{x_k}).$$

Denote $O_j = \bigcup_{k=1}^{j} B(x, \varepsilon_{x_k})$. Clearly, $O_j \subseteq O$ is an open bounded set and $\{O_j\}\uparrow O$. $\square$

6.13 Let (X, τ) be a Hausdorff space and C and K be disjoint compact sets. Show that there are disjoint open supersets, U and V, of C and K, respectively.

Solution. If one of the compact sets C or K is empty, then the proof is trivial. Let $C = \{x\}$ be a singleton. Because X is Hausdorff, for each $y \in K$, there is a pair U_x and V_y of disjoint open neighborhoods of x and y. Now, for all y's of K, we have thereby two systems $\{U_{x_k}, V_{y_k}\}$ of open sets, of which $\{V_{y_k}\}$ is an open cover of K and that due to compactness of K can be reduced to a finite subcover, say $V_{y_1}, \ldots, V_{y_n}$. Hence the sets $U = \bigcap_{k=1}^{n} U_{x_k}$ and $V = \bigcup_{k=1}^{n} V_{y_k}$ are those with the desired property. Suppose now that C has more than one point. Then, for each $x \in C$, as we showed above, there is an open

neighborhood O_x of x and an open superset W_x of K, which are disjoint. For all $x \in C$, we thereby form two systems, $\{O_{x_m}, W_{y_m}\}$, of which $\{O_{x_m}\}$ is an open cover of C. This cover can be reduced to a finite subcover, $O_{x_1}, \ldots, O_{x_s}$. Hence the sets

$$U = \bigcup_{m=1}^{s} O_{x_m} \quad \text{and} \quad V = \bigcap_{m=1}^{s} W_{y_m}$$

are of the required property. □

6.15 Let $\mathcal{K}$ be the family of all compact subsets of a topological space (X, τ). Under the formation of which set operations is $\mathcal{K}$ closed? Give counterexamples of those set operations, which are not contained in $\mathcal{K}$.

Solution. Obviously $\mathcal{K}$ is not closed under the formation of differences and complements. Let $A, B \in \mathcal{K}$ and let $\{O_i\}$ be an open cover of $A \cup B$. Then, $\{O_i\}$ is a cover both either A and B. Because A and B are compact, the cover $\{O_i\}$ is reducible to two finite subcovers whose union is a finite cover of $A \cup B$. Thus, $A \cup B$ is compact. The same applies to any finitely many compact sets. An infinite union of compact sets need not be compact. For example, let $K_n = \left[\frac{1}{n}, 1 - \frac{1}{n}\right]$ in $(\mathbb{R}, d_e)$. The union $\bigcup_{n=1}^{\infty} K_n = (0, 1)$ is not compact.

A compact set in a topological space need not be closed unless it is Hausdorff. If $\{K_i; i \in I\} \subseteq \mathcal{K}$ is an arbitrary indexed family in a Hausdorff space, then, the intersection $\bigcap_{i \in I} K_i$ is closed and is a subset of any one of K_i. Thus, by Theorem 6.9, it is compact. However, it does not hold true in general topological spaces.

In conclusion, $\mathcal{K}$ is closed with respect to the formation of finite unions and, if τ is Hausdorff, under arbitrary intersections. Furthermore, the empty set is also compact. □

Chapter 4. Measurable Spaces and Measurable Functions

1.2 Let $\mathcal{R}$ be a ring in Ω. Show that with any two sets $A, B \in \mathcal{R}$, their intersection also belongs to $\mathcal{R}$.

Solution. The statement follows from

$$A \cap B = (A \cup B) \backslash (A \Delta B),$$

where

$$A \Delta B = (A \backslash B) \cup (B \backslash A). \qquad \square$$

1.4 Let $\mathcal{A}$ be an algebra in Ω. Prove that $\mathcal{A}$ is a σ-algebra in Ω if and only if $\mathcal{A}$ is also a monotone system.

Solution. Obviously, any σ-algebra is a monotone system. Let $\mathcal{A}$ be an algebra and monotone system. Let $\{A_n\} \subseteq \mathcal{A}$. We show that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Let $B_N = \bigcup_{n=1}^{N} A_n$. Then, $B_N \in \mathcal{A}$ and because $\mathcal{A}$ is a monotone system and $\{B_N\}$ is monotone nondecreasing,

$$\bigcup_{N=1}^{\infty} B_N \in \mathcal{A}. \qquad \square$$

1.6 Give an infinite collection of subsets of $\mathbb{R}$ that contains $\emptyset$ and $\mathbb{R}$ and which is closed under countable unions and intersections but is not a σ-algebra.

Solution. Let $\mathcal{S} = \{\mathbb{R}, \emptyset, A_n : n \in \mathbb{N}\}$, where $A_n = (-\infty, n)$. Clearly, $\mathcal{S}$ meets the assumptions of the problem. However, $\mathcal{S}$ does not contain complements. For example, $A_n^c = [n, \infty) \notin \mathcal{S}$. Thus, $\mathcal{S}$ is not even an algebra, but it is a monotone system. $\square$

1.8 Give a Dynkin system that is not a σ-algebra.

Solution. The system in Problem 1.7 is one, because the intersection of two overlapping sets does not have to contain an even number of elements. □

1.10 Prove the statement: *A Dynkin system $\mathcal{D}$ is a σ-algebra if and only if $\mathcal{D}$ is $\cap$-stable.*

Solution. (i) Assume that $\mathcal{D}$ is a $\cap$-stable Dynkin system. Let $A, B \in \mathcal{D}$. Because $\mathcal{D}$ is $\cap$-stable $\Rightarrow A \cap B \in \mathcal{D}$. Furthermore, because $A \cap B \subseteq A$ we can apply the result of Problem 1.9, namely,

$$A \backslash B = A \backslash (A \cap B) \in \mathcal{D}.$$

Therefore, an $\cap$-stable Dynkin system is closed relative to differences. Next, because $A \cup B = A + (B \backslash A)$, it follows that a $\cap$-stable Dynkin system is closed relative to finite unions.

Now, let $\{A_n\} \subseteq \mathcal{D}$. Denote $B_n = \bigcup_{k=1}^{n} A_k, n = 1, 2, \ldots$, and $B_0 = \emptyset$. Obviously, $\{B_n\} \subseteq \mathcal{D}$ and sequence $\{B_n\}$ is monotone non-decreasing. Then,

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=0}^{\infty} B_n = \sum_{n=1}^{\infty} C_n, \tag{P1.10}$$

where $C_n = B_n \backslash B_{n-1}, n = 1, 2, \ldots$ are disjoint. Moreover, $C_n \in \mathcal{D}$, again by Problem 1.9. Thus, from (P1.10) we conclude that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

(ii) Let $\mathcal{D}$ be a σ-algebra. Then, by Problem 1.5, it is also a Dynkin system. As a σ-algebra, $\mathcal{D}$ is $\cap$-stable (based on Problems 1.2 and 1.5 or Example 1.2(v). □

1.12 Show the equivalence of two definitions of a semi-ring if property c) in Definition 1.1(v) can be replaced with

c') Let $A, A_1, \ldots, A_n \in \mathcal{S}$. Then there is a finite tuple $C_1, \ldots, C_k$ of disjoint sets from $\mathcal{S}$ such that

$$A \setminus \bigcup_{s=1}^{n} A_s = \sum_{j=1}^{k} C_j.$$

Solution. Let $\mathcal{S}$ be a semi-ring in the sense of Definition 1.1(v). Then,

$$\begin{aligned} A\setminus(A_1 \cup A_2) &= (A\setminus A_1)\setminus A_2 \\ &= \{\textstyle\sum_{j=1}^{r} C_j\}\setminus A_2 = \sum_{j=1}^{r} C_j \setminus A_2 \\ &= \textstyle\sum_{j=1}^{r}\sum_{k=1}^{n_j} D_{k_j}. \end{aligned}$$

The statement follows by induction. The converse is trivial. □

2.1 Prove Theorem 2.1. *The intersection of arbitrarily many* σ*-algebras* (*algebras*, *monotone systems*, *rings*) *in* Ω *is a* σ*-algebra* (*an algebra*, *a monotone system*, *a ring*).

Solution. Let $\Sigma = \bigcap_{i\in I} \Sigma_i$, where $\{\Sigma_i : i \in I\}$ is an indexed family of σ-algebras in Ω. Then,

(a) $\Omega \in \Sigma$.

(b) If $A \in \Sigma$, then A belongs to all Σ_i's and thus $A^c \in \Sigma_i, i \in I$. Consequently, $A^c \in \Sigma$.

(c) Let $\{A_n\} \subseteq \Sigma$. Then, each $A_n \in \Sigma_i$ for all $i \in I$ and hence, $\bigcup_{n=1}^{\infty} A_n \in \Sigma_i$ for all i's implying that the union of A_n's is also an element of Σ.

The proof regarding the intersection of algebras, monotone systems, and rings is similar. □

2.3 Show that a union of σ-algebras in Ω need not be a σ-algebra.

Solution. Let $\Sigma_1 = \{\Omega, \emptyset, A, A^c\}$ and $\Sigma_2 = \{\Omega, \emptyset, B, B^c\}$ and let $A \neq B$. Then, $\Sigma_1 \cup \Sigma_2 = \{\Omega, \emptyset, A, A^c, B, B^c\}$. Excluding some trivial cases, $\Sigma_1 \cup \Sigma_2$ is not a σ-algebra. □

2.5 Let Σ be a σ-algebra in Ω and let $B \subseteq \Omega$. Show that the σ-algebra generated by $\mathcal{G} = \Sigma \cup \{B\}$ is of the form

$$\Sigma' = \{(A_1 \cap B) \cup (A_2 \backslash B) \colon\ A_1, A_2 \in \Sigma\}.$$

[*Hint*: 1) Show that $\sigma(\Sigma') = \sigma(\Sigma \cup \{B\})$. 2) Show that Σ' is a σ-algebra in Ω.]

Solution. Denote $\sigma(\Sigma')$ the σ-algebra generated by Σ' and $\sigma(\Sigma \cup \{B\})$ the σ-algebra generated by $\Sigma \cup \{B\}$. We prove the statements in accordance with the following chronology.

$$1)\ \ \sigma(\Sigma') = \sigma(\Sigma \cup \{B\}); \quad 2)\ \ \Sigma' \text{ is a } \sigma\text{-algebra.}$$

1) We show that $\sigma(\Sigma') = \sigma(\Sigma \cup \{B\})$

(i) Notice first that B need not be a measurable set. Thus, the smallest σ-algebra generated by Σ and $\{B\}$ must contain all sets of type $A_1 \cap B$, $A_2 \backslash B$, and $(A_1 \cap B) \cup (A_2 \backslash B)$ such that $A_1, A_2 \in \Sigma$, to say the least. Hence, $\Sigma' \subseteq \sigma(\Sigma \cup \{B\})$ implying that

$$\sigma(\Sigma') \subseteq \sigma(\Sigma \cup \{B\}), \tag{P2.5}$$

because $\sigma(\Sigma')$ is the smallest σ-algebra containing Σ' and $\sigma(\Sigma \cup \{B\})$ is a σ-algebra containing Σ'.

(ii) We show the inverse inclusion of (P2.5). We list the following properties of Σ' as per its definition.

$a)$ With the choice of $A_1 = \Omega$ and $A_2 = \emptyset$, we have $B \in \Sigma'$.

$b)$ If $A_1 = A_2 = A$, then $(A_1 \cap B) \cup (A_2 \backslash B) = A \in \Sigma'$ implying that $\Sigma \subseteq \Sigma'$.

Therefore,

$$\Sigma \cup \{B\} \subseteq \Sigma' \ \Rightarrow \sigma(\Sigma \cup \{B\}) \subseteq \sigma(\Sigma'),$$

consequently having

$$\sigma(\Sigma') = \sigma(\Sigma \cup \{B\})$$

and we are done with $1)$.

$2)$ We show that Σ' is a σ-algebra.

(a) Because $\Sigma \subseteq \Sigma'$, it follows that $\Omega \in \Sigma'$.

(b) Let $C \in \Sigma'$. Then,

$$C = (A_1 \cap B) \cup (A_2 \backslash B) = (A_1 \cap B) \cup (A_2 \cap B^c).$$

By DeMorgan's and distributive laws,

$$C^c = (A_1^c \cap A_2^c) \cup (A_2^c \cap B^c) \cup (A_1^c \cap B) \cup (B \cap B^c).$$

Now, because $A_1^c \cap A_2^c \subseteq (A_2^c \cap B^c) \cup (A_1^c \cap B)$, which can be shown by a pick-a-point process, we have C^c reduce to

$$C^c = (A_2^c \cap B^c) \cup (A_1^c \cap B) = (A_1^c \cap B) \cup (A_2^c \backslash B) \in \Sigma',$$

in as much as $A_1^c, A_2^c \in \Sigma$.

(c) Let $\{C_n\} \subseteq \Sigma'$. Then, $C_n = (A_1^n \cap B) \cup (A_2^n \backslash B)$ and

$$\bigcup_{n=1}^{\infty} C_n = \left\{ \bigcup_{n=1}^{\infty} (A_1^n \cap B) \right\} \cup \left\{ \bigcup_{n=1}^{\infty} (A_2^n \cap B^c) \right\}$$

$$= \left\{ \underbrace{\left(\bigcup_{n=1}^{\infty} A_1^n \right)}_{\in \Sigma} \cap B \right\} \cup \left\{ \underbrace{\left(\bigcup_{n=1}^{\infty} A_2^n \right)}_{\in \Sigma} \backslash B \right\} \in \Sigma'.$$

In conclusion, Σ' is a σ-algebra. $\square$

2.7 Let Ω be an arbitrary nonempty set and let $A, B \subseteq \Omega$. Construct for the topology $\tau = \{\Omega, \emptyset, A, B, A \cap B, A \cup B\}$ the Borel σ-algebra $\mathcal{B}(\tau)$.

Solution. Supplement τ with the following subsets.

$$A\backslash B, B\backslash A, A^c, B^c, A \cup B^c, A^c \cup B, A^c \cap B^c, A^c \cup B^c, A\Delta B,$$

and $(A\Delta B)^c = (A^c \cup B) \cap (A \cup B^c)$. □

2.9 Show that any open set in $\mathbb{R}^n$ can be represented as an at most countable union of disjoint half-open cubes.

Solution. For each positive integer k, let $\mathfrak{P}_k$ be the partition of $\mathbb{R}^n$ by cubes of type $[a, b)$, where

$$a_i = \tfrac{m}{2^k},\ b_i = \tfrac{m+1}{2^k},\ m = 0, \pm 1, \pm 2, \ldots$$

For $k < p$, clearly $\mathfrak{P}_k \prec \mathfrak{P}_p$ in the sense that each cube of $\mathfrak{P}_k$ is included in some cube of $\mathfrak{P}_p$.

Let G be a non-empty open set in $\mathbb{R}^n$. Let S_1 be the set of all cubes from $\mathfrak{P}_1$ contained in G. If $S_1 = \emptyset$, we try the same with $\mathfrak{P}_2$ and so on. Suppose k is the smallest positive integer for which $S_k \neq \emptyset$. Denote S_k by σ_k. From S_{k+1} select only those cubes that are not included in σ_k and denote this set by σ_{k+1} and continue this process to form the sequence $\{\sigma_j\}$. Now let Σ_j stand for the union of all cubes from σ_j. Clearly, $\Sigma_j \subseteq G$ and therefore

$$G \supseteq \bigcup_{j=k}^{\infty} \Sigma_j = \Sigma.$$

We claim that

$$G = \bigcup_{j=k}^{\infty} \Sigma_j$$

Let $x \in G$. Because G is open, there is an open ball $B(x, r) \subseteq G$. By satisfying the inequality

$$\frac{\sqrt{n}}{2^t} < r$$

with respect to t we can select exactly one cube, say $C(x,t)$, from $\mathfrak{P}_t$, which contains point x and is included in $B(x,r)$ and therefore in G. Surely, the inequality will also be satisfied for $t+1$ with one cube available from $\mathfrak{P}_{t+1}$ which is in G and so on. Each one such cube $C(x,t)$ is from σ_t and thus is included in Σ_t. The latter contains x and therefore $x \in \Sigma$. □

3.1 Show that $\widetilde{\Sigma}$ in the proof of Proposition 3.4 is a σ-algebra.

Solution.

(i) Obviously, $\Omega' \in \widetilde{\Sigma}$.

(ii) Let $\widetilde{A} \in \widetilde{\Sigma}$. Then, $f^*\left(\widetilde{A}^c\right) = \left(f^*\left(\widetilde{A}\right)\right)^c \in \Sigma$, because $f^*\left(\widetilde{A}\right)$ does. Hence $\widetilde{A}^c \in \widetilde{\Sigma}$.

(iii) Next, if $\left\{\widetilde{A}_n\right\} \subseteq \widetilde{\Sigma} \Rightarrow f^*\left(\widetilde{A}_n\right) \in \Sigma$. Then,

$$f^*\left(\bigcup_{n=1}^{\infty} \widetilde{A}_n\right) = \bigcup_{n=1}^{\infty} f^*\left(\widetilde{A}_n\right) \in \Sigma,$$

implying that $\widetilde{\Sigma}$ is closed relative to countable unions. □

3.3 Let $f : \Omega \to \Omega'$ be a function and let $\mathcal{G}' \subseteq \mathcal{P}(\Omega')$. Show that

$$f^{**}(\Sigma(\mathcal{G}')) = \Sigma(f^{**}(\mathcal{G}')).$$

[*Hint*: Let $\widetilde{\Sigma} := \{A' \in \mathcal{P}(\Omega') : f^*(A') \in \Sigma(f^{**}(\mathcal{G}'))\}$. Show that $\widetilde{\Sigma}$ is a σ-algebra. Then show that $\Sigma(\mathcal{G}') \subseteq \widetilde{\Sigma}$.]

Solution. (i) Because $\mathcal{G}' \subseteq \Sigma(\mathcal{G}')$, we have $f^{**}(\mathcal{G}') \subseteq f^{**}(\Sigma(\mathcal{G}'))$. Obviously, $f^{**}(\Sigma(\mathcal{G}'))$ is a σ-algebra in Ω containing $f^{**}(\mathcal{G}')$. Because $\Sigma(f^{**}(\mathcal{G}'))$ is the smallest σ-algebra generated by $f^{**}(\mathcal{G}')$,

$$\Sigma(f^{**}(\mathcal{G}')) \subseteq f^{**}(\Sigma(\mathcal{G}')).$$

(ii) We show the inverse inclusion. Define

$$\widetilde{\Sigma} := \{A' \in \mathcal{P}(\Omega') : f^*(A') \in \Sigma(f^{**}(\mathcal{G}'))\}. \tag{P3.3}$$

We show that $\widetilde{\Sigma}$ is a σ-algebra in Ω':

$a)$ $f^*(\Omega') = \Omega \in \Sigma(f^{**}(\mathcal{G}')) \Rightarrow \Omega' \in \widetilde{\Sigma}$.

$b)$ Let $A' \in \widetilde{\Sigma}$. Then, $f^*(A') \in \Sigma(f^{**}(\mathcal{G}'))$ and $f^*(A'^c) = [f^*(A')]^c \in \Sigma \Rightarrow A'^c \in \widetilde{\Sigma}$.

$c)$ If $\{A'_n\} \subseteq \Sigma \Rightarrow \{f^*(A'_n)\} \subseteq \Sigma(f^{**}(\mathcal{G}')) \Rightarrow \bigcup_{n=1}^{\infty} f^*(A'_n) \in \Sigma(f^{**}(\mathcal{G}')) \Rightarrow f^*\left(\bigcup_{n=1}^{\infty} A'_n\right) \in \Sigma(f^{**}(\mathcal{G}')) \Rightarrow \bigcup_{n=1}^{\infty} A'_n \in \widetilde{\Sigma}$.

(iii) We show that $\mathcal{G}' \subseteq \widetilde{\Sigma}$. Obviously, $\forall G' \in \mathcal{G}', f^*(G') \in f^{**}(\mathcal{G}')$ bearing in mind that

$$f^{**}(\mathcal{G}') = \{f^*(G') : G' \in \mathcal{G}'\}.$$

Consequently,

$$f^*(G') \in \Sigma(f^{**}(\mathcal{G}')) \text{ good } \forall G' \in \mathcal{G}' \Rightarrow G' \in \widetilde{\Sigma}.$$

(iv) From (iii) it follows that $\Sigma(\mathcal{G}') \subseteq \widetilde{\Sigma}$ and thus,

$$f^{**}(\Sigma(\mathcal{G}')) \subseteq f^{**}\big(\big(\widetilde{\Sigma}\big)\big) \subseteq \Sigma(f^{**}(\mathcal{G}')).$$

The second inclusion is the direct consequence of the definition of $\widetilde{\Sigma}$ in (P3.3).

So, the dual containment yields the asserted equation. □

3.4 Let $[O, O_1, F]$ be a homeomorphism, with O and O_1 being open sets in topological spaces (X, τ) and (X, τ_1), respectively, and let $\mathcal{B}(\tau_O)$ and $\mathcal{B}(\tau_{O_1})$ be the Borel σ-algebras generated

by the relative topologies τ_O and τ_{O_1}. Prove that $[\mathcal{B}(\tau_O), \mathcal{B}(\tau_{O_1}), F_*]$ is bijective.

Solution. We prove that $F_{**}(\mathcal{B}_O) = \mathcal{B}_{O_1}$ and $F^{**}(\mathcal{B}_{O_1}) = \mathcal{B}_O$. If τ_O and τ_{O_1} are the respective relative topologies, by Problem 3.11, Chapter 1,

$$F_{**}(\tau_O) = \tau_{O_1} \text{ and } F^{**}(\tau_{O_1}) = \tau_O. \tag{P3.4}$$

By notation,

$$\sigma(\tau_{O_1}) = \sigma(O_1 \cap \tau_1) = \mathcal{B}_{O_1}. \tag{P3.4a}$$

By Problem 3.3, and (P3.4 - P3.4a),

$$F^{**}(\mathcal{B}_{O_1}) = F^{**}(\sigma(\tau_{O_1})) = \sigma(F^{**}(\tau_{O_1}))$$

$$= \sigma(\tau_O) = \mathcal{B}_O.$$

Because F is homeomorphism, applying the same procedure to F_{**}, we prove that $F_{**}(\tau_O) = \tau_{O_1}$. □

Chapter 5. Measures

1.2 Prove Proposition $1.5(i, ii)$. *Every content μ on a ring $\mathcal{R}$ has the following properties*.

(i) $A \subseteq B \Rightarrow \mu(B \backslash A) = \mu(B) - \mu(A)$ [provided that $\mu(A) < \infty$].

(ii) $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

Solution.

(i) $A \subseteq B \Rightarrow B = A + (B \backslash A)$. Because μ is a content, thus additive,

$$\mu(B) = \mu(A) + \mu(B\backslash A).$$

Because $\mu(B\backslash A) \geq 0 \Rightarrow \mu(B) \geq \mu(A)$. Now, if $\mu(A) < \infty$, we have

$$\mu(B) - \mu(A) = \mu(B\backslash A).$$

(ii) $A \cup B = A + (B\backslash A) = A + (B\backslash(A \cap B))$. Then by (i) and additivity of a content it follows that

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B). \qquad \square$$

1.4 Prove that μ is a measure on Σ in Example $1.2(v)$.

Solution. Let $B_n = \bigcup_{k=n}^{\infty} A_n$. Then, $\{B_n\}$ is monotone decreasing and $\mu(B_1) \leq \sum_{n=1}^{\infty} \mu(A_n) < \infty$. Thus, we can apply the continuity from above of Proposition 1.5 (ii) to the sequence $\{B_n\}$ with

$$B = \bigcap_{n=1}^{\infty} B_n = \overline{\lim_{n \to \infty}} A_n$$

so that

$$\lim_{n \to \infty} \mu(B_n) = \mu(B) = \mu\big(\overline{\lim_{n \to \infty}} A_n\big).$$

On the other hand, $\mu(B_n) \leq \sum_{k=n}^{\infty} \mu(A_k) \downarrow 0$ as $n \to \infty$ [due to the convergence of $\sum_{n=1}^{\infty} \mu(A_n)$]. Therefore, $\mu(B) = 0$. $\square$

1.6 A subset $\mathcal{R}_I$ of a ring $\mathcal{R}$ in Ω is called an *ideal in* $\mathcal{R}$ if it has the following properties.

$a)$ $\emptyset \in \mathcal{R}_I$.

$b)$ $B \in \mathcal{R}_I$, $A \subseteq B$, $A \in \mathcal{R} \Rightarrow A \in \mathcal{R}_I$.

$c)$ A, $B \in \mathcal{R}_I \Rightarrow A \cup B \in \mathcal{R}_I$.

Let μ be a content on a ring $\mathcal{R}$. Define $\mathcal{R}_\mu = \{R \in \mathcal{R} : \mu(R) = 0\}$. Show that $\mathcal{R}_\mu$ is an ideal in $\mathcal{R}$.

Solution. (i) Because $\mu(\emptyset) = 0 \Rightarrow \emptyset \in \mathcal{R}_\mu$.

(ii) Let $B \in \mathcal{R}_\mu$. Then $\mu(B) = 0$. Let $A \in \mathcal{R}$ such that $A \subseteq B$ (at least one such subset of $\mathcal{R}$ clearly exists). Then, because $\mu(A) \leq \mu(B)$, $\mu(A) = 0$ and thus $A \in \mathcal{R}_\mu$.

(iii) Let $A, B \in \mathcal{R}_\mu$. Then,

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) = 0 + 0 - 0 = 0.$$

Thus, $A \cup B \in \mathcal{R}_\mu$. □

1.8 Let $\mathcal{R}_{\sigma\text{-}I}$ be a σ-ideal in a ring $\mathcal{R}$. Show that there exists a premeasure μ on $\mathcal{R}$ such that $\mathcal{R}_{\sigma\text{-}I} = \mathcal{R}_\mu$, defined in Problems 1.7 and 1.8.

Solution. Define

$$\mu(A) = \begin{cases} 0, & A \in \mathcal{R}_{\sigma\text{-}I} \\ \infty, & A \in \mathcal{R} \backslash \mathcal{R}_{\sigma\text{-}I}. \end{cases}$$

We show that μ is a premeasure on $\mathcal{R}$.

$a)$ $\mu(\emptyset) = 0$ and $\mu \geq 0$.

$b)$ Let $\{A_n\} \subseteq \mathcal{R}$ be disjoint such that $\sum_{n=1}^{\infty} A_n \in \mathcal{R}$.

Case 1: $\{A_n\} \subseteq \mathcal{R}_{\sigma\text{-}I}$. Then, by Problem 1.7, $\sum_{n=1}^{\infty} A_n \in \mathcal{R}_{\sigma\text{-}I}$ implying that $\mu(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n) = 0$.

Case 2: Let $J \subseteq \mathbb{N}$ be such that $\{A_{n_k} : n_k \in J\} \subseteq \mathcal{R} \backslash \mathcal{R}_{\sigma\text{-}I}$. Then,

$$\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n \in J} \mu(A_n) + \sum_{n \in \mathbb{N} \backslash J} \mu(A_n) = \infty + 0 = \infty.$$

On the other hand, $\sum_{n=1}^{\infty} A_n \notin \mathcal{R}_{\sigma\text{-}I}$, because if it were the case, by the definition of an ideal, those A_n with $n \in J$, as subsets of

$\sum_{n=1}^{\infty} A_n$ would have been elements of $\mathcal{R}_{\sigma\text{-}I}$ and the latter was not assumed. Consequently, $\mu(\sum_{n=1}^{\infty} A_n) = \infty$ and hence

$$\mu(\textstyle\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

Observe that $\mathcal{R}_{\sigma\text{-}I}$ contains only elements of $\mathcal{R}$ whose measure is zero. In short, $\mathcal{R}_{\sigma\text{-}I} = \{A \in \mathcal{R} : \mu(A) = 0\}$. □

1.10 Let (Ω, Σ, μ) be a measure space and $\{A_n\} \subseteq \Sigma$ be a sequence of sets with $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) < \infty$ and $\inf\{\mu(A_n)$: $n = 1, 2, \dots\} = a \geq 0$. Show that $\mu\big(\varlimsup_{n\to\infty} A_n\big) \geq a$.

Solution. Let $B_n = \bigcup_{k=n}^{\infty} A_k$. The sequence $\{B_n\}$ is monotone non-increasing with $\mu(B_1) < \infty$. Therefore, by continuity from above [Proposition 1.5(ii)] and because obviously $\mu(B_n) \geq a$ for all n,

$$\varlimsup_{n\to\infty} \mu(B_n) = \mu\left(\bigcap_{n=1}^{\infty} B_n\right) = \mu\big(\varlimsup_{n\to\infty} A_n\big) \geq a. \qquad \square$$

1.12 Under the condition of Problem 1.11, let $a = \infty$ and μ be a finite measure. Show that $\Sigma_\infty = \Sigma$.

Solution. Obviously, $\mathcal{G}_\infty = \Sigma$ and therefore $\Sigma_\infty = \{Q \subseteq \Omega$: $Q \cap A \in \Sigma,\ \forall A \in \Sigma\}$. Because $\Sigma \subseteq \Sigma_\infty$, it remains to show the inverse inclusion. Let $A_\infty \in \Sigma_\infty$. Then, for every $A \in \Sigma$, $A_\infty \cap A$, $A_\infty \cap A^c$, and, therefore, their union,

$$A_\infty = (A_\infty \cap A) + (A_\infty \cap A^c),$$

are elements of Σ. This yields that $\Sigma_\infty \subseteq \Sigma$. □

1.14 Argue that for any probability space (Ω, Σ, μ), the axiom $\mu(\emptyset) = 0$ is redundant. Is it also true for any measure?

Solution. From

$$\mu(\emptyset + \emptyset) = \mu(\emptyset) + \mu(\emptyset) = \mu(\emptyset) \tag{P1.14}$$

(due to the additivity of μ) it follows that $\mu(\emptyset) = 0$. Thus the latter as an axiom of a measure is redundant. However, we managed to prove it because μ is a finite measure. In the general case, if μ is not finite the equation $\mu(\emptyset) = 0$ cannot be concluded from (P1.14). □

1.16 Prove Lemma 1.11. *Let* $\mu_1, \mu_2 \in \mathfrak{M}(\Omega, \Sigma)$. *Then the set function*

$$\nu(A) = \inf\{\mu_1(E) + \mu_2(A\backslash E) : E \in \Sigma \cap A\}$$

is a measure on Σ.

Solution. Obviously, ν is a nonnegative set function defined on Σ that equals zero at the empty set.

Next we show that ν is monotone. Let $E \subseteq A \subseteq B$. Then, we have

$$\mu_1(E) + \mu_2(A\backslash E) \leq \mu_1(E) + \mu_2(B\backslash E).$$

Because this inequality holds true for all $E \subseteq A$,

$$\nu(A) \leq \mu_1(E) + \mu_2(B\backslash E)$$

and because $E \subseteq B$, we have $\nu(A) \leq \nu(B)$.

We need to prove that ν is σ-additive. Let $\{A_n\} \subseteq \Sigma$ be a sequence of disjoint sets with $A := \sum_{n=1}^{\infty} A_n$. We first show that

$$\nu(A) \geq \sum_{n=1}^{\infty} \nu(A_n). \tag{P1.16}$$

Let E be a measurable subset of A. Then,

$$\begin{aligned} &\mu_1(E) + \mu_2(A\backslash E) \\ &= \mu_1\left(\sum_{n=1}^{\infty}(A_n \cap E)\right) + \mu_2\left(\sum_{n=1}^{\infty}(A_n \cap E^c)\right) \\ &= \sum_{n=1}^{\infty} \mu_1(A_n \cap E) + \mu_2(A_n\backslash(A_n \cap E)) \end{aligned}$$

$$\geq \sum_{n=1}^{\infty} \nu(A_n).$$

Because $\sum_{n=1}^{\infty} \nu(A_n) \leq \mu_1(E) + \mu_2(A \backslash E)$ holds for all subsets E of A, then it also holds for

$$\inf\{\mu_1(E) + \mu_2(A \backslash E) : E \in \Sigma \cap A\} = \nu(A)$$

and thereby (P1.16) is proved.

It remains to show the inverse inequality. Let $\varepsilon > 0$ be chosen. From the definition of ν, given A_n, and any $\varepsilon_n > 0$, there is a subset $E_n \subseteq A_n$ such that

$$\nu(A_n) > \mu_1(E_n) + \mu_2(A_n \backslash E_n) - \varepsilon_n,\ n = 1, 2, \ldots .$$

We choose ε_n's such that $\sum_{n=1}^{\infty} \varepsilon_n \leq \varepsilon$. Denote $E = \sum_{n=1}^{\infty} E_n$. Obviously, E_n's are disjoint and $E \subseteq A$ and

$$A \backslash E = \sum_{n=1}^{\infty} (A_n \backslash \sum_{k=1}^{\infty} E_k) = \sum_{n=1}^{\infty} (A_n \backslash E_n),$$

because the rest of E_k's are disjoint with A_n. Therefore, with

$$\mu_1(E) + \mu_2(A \backslash E)$$

$$= \sum_{n=1}^{\infty} \mu_1(E_n) + \sum_{n=1}^{\infty} \mu_2(A_n \backslash E_n)$$

we have:

$$\sum_{n=1}^{\infty} \nu(A_n) \geq \sum_{n=1}^{\infty} \mu_1(E_n)$$

$$+ \sum_{n=1}^{\infty} \mu_2(A_n \backslash E_n) - \sum_{n=1}^{\infty} \varepsilon_n$$

$$\geq \mu_1(E) + \mu_2(A \backslash E) - \varepsilon \geq \nu(A) - \varepsilon.$$

Because this inequality holds for each $\varepsilon > 0$, we arrive at the desired inverse inequality and we are done with the lemma. □

1.18 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and let $\mathcal{G} \subseteq \Sigma$ be a family of events. $\mathcal{G}$ is called *independent* if for any finite tuple of

events $G_1, \ldots, G_m$, $\mathbb{P}\left(\bigcap_{i=1}^{m} G_i\right) = \prod_{i=1}^{m} \mathbb{P}(G_i)$. Let $\mathcal{D}(\mathcal{G})$ be the Dynkin system induced by $\mathcal{G}$. Show that $\mathcal{D}(\mathcal{G})$ is also independent.

Solution. Define the system of events (measurable sets) as follows.

$$\widetilde{\mathcal{D}} = \{A \subseteq \Sigma : \{A\} \cup \mathcal{G} \text{ is independent}\}.$$

We show that $\widetilde{\mathcal{D}}$ is Dynkin. Obviously, $\Omega \in \widetilde{\mathcal{D}}$. Let some $A \in \widetilde{\mathcal{D}}$ and let $G_1, \ldots, G_k \in \mathcal{G}$. Denote $G = \bigcap_{i=1}^{k} G_i$. Then, events A and G are independent. Hence,

$$\mathbb{P}(A \cap G) = \mathbb{P}(A \cap G) = \mathbb{P}(A)\mathbb{P}(G)$$

$$= \mathbb{P}(A) \prod_{i=1}^{k} \mathbb{P}(G_i).$$

Now, $A^c \cap G = G \backslash (A \cap G)$ implies that

$$\mathbb{P}(A^c \cap G) = \mathbb{P}(G) - \mathbb{P}(A)\mathbb{P}(G) = [1 - \mathbb{P}(A)]\mathbb{P}(G)$$

$$\mathbb{P}(A^c)\mathbb{P}(G).$$

Thus, $A^c, G_1, \ldots, G_k$ are independent and $A^c \in \widetilde{\mathcal{D}}$. Now, let $\{A_n\}$ be a sequence of exclusive (disjoint) events from $\widetilde{\mathcal{D}}$. We show that $\sum_{n=1}^{\infty} A_n \in \widetilde{\mathcal{D}}$. Let again $G_1, \ldots, G_k \in \mathcal{G}$ and $G = \bigcap_{i=1}^{k} G_i$. Then, events A_n and G are independent, that is, for each A_n,

$$\mathbb{P}(A_n \cap G) = \mathbb{P}(A_n \cap G)$$

$$= \mathbb{P}(A_n)\mathbb{P}(G) = \mathbb{P}(A_n) \prod_{i=1}^{k} \mathbb{P}(G_i).$$

Consider

$$\mathbb{P}(G \cap \sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(G \cap A_n)$$

$$= \sum_{n=1}^{\infty} \mathbb{P}(G)\mathbb{P}(A_n)$$

$$= \mathbb{P}(G)\mathbb{P}(\sum_{n=1}^{\infty} A_n).$$

Hence, $\widetilde{\mathcal{D}}$ indeed is Dynkin. Now, because $\mathcal{G} \subseteq \widetilde{\mathcal{D}}$, then $\mathcal{D}(\mathcal{G}) \subseteq \widetilde{\mathcal{D}}$ and therefore, $\mathcal{D}(\mathcal{G})$ as a subsystem of an independent family, is independent. □

1.20 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and let $\{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ be a set of n families of events. This set is said to be *independent* if for any choice of events $G_1 \in \mathcal{G}_1, \ldots, G_n \in \mathcal{G}_n$, the tuple $\{G_1, \ldots, G_n\}$ is independent. If the set $\{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ is independent, show that the set $\{\mathcal{D}(\mathcal{G}_1), \ldots, \mathcal{D}(\mathcal{G}_n)\}$ (of Dynkin systems induced by $\mathcal{G}_i$'s) is also independent.

Solution. Define the following family of events:

$$\widetilde{\mathcal{D}}_1 = \{A \subseteq \Sigma : \{\{A\} \cup \mathcal{G}_1, \ldots, \mathcal{G}_n\} \text{ is independent}\}.$$

Here $\mathcal{G}_1$ is enhanced, whereas the rest of $\mathcal{G}_i$'s are the same. We show that $\widetilde{\mathcal{D}}_1$ is Dynkin. Obviously, $\Omega \in \widetilde{\mathcal{D}}_1$. Let some $A \in \widetilde{\mathcal{D}}_1$ and let $G_2 \in \mathcal{G}_2, \ldots, G_n \in \mathcal{G}_n$. (Note that A can also belong to $\mathcal{G}_1$.) Denote $G = \bigcap_{i=2}^{n} G_i$. Then, events A and G are independent, implying that

$$\mathbb{P}(A \cap G) = \mathbb{P}(A \cap G) = \mathbb{P}(A)\mathbb{P}(G) = \mathbb{P}(A)\prod_{i=2}^{n} \mathbb{P}(G_i).$$

Now, $A^c \cap G = G \backslash (A \cap G)$ yields

$$\mathbb{P}(A^c \cap G) = \mathbb{P}(G) - \mathbb{P}(A)\mathbb{P}(G)$$

$$= [1 - \mathbb{P}(A)]\mathbb{P}(G)$$

$$= \mathbb{P}(A^c)\mathbb{P}(G).$$

Thus, $A^c, G_2, \ldots, G_n$ are independent and $A^c \in \widetilde{\mathcal{D}}_1$. Now, let $\{A_r\}$ be a sequence of exclusive (disjoint) events from $\widetilde{\mathcal{D}}_1$. We show that $\sum_{r=1}^{\infty} A_r \in \widetilde{\mathcal{D}}_1$. Let again $G_2 \in \mathcal{G}_2, \ldots, G_n \in \mathcal{G}_n$ and $G = \bigcap_{i=2}^{n} G_i$. Then, events A_r and G are independent, that is, for each A_r,

$$\mathbb{P}(A_r \cap G) = \mathbb{P}(A_r \cap G)$$

$$= \mathbb{P}(A_r)\mathbb{P}(G) = \mathbb{P}(A_r) \prod_{i=2}^{n} \mathbb{P}(G_i).$$

Consider

$$\mathbb{P}(G \cap \textstyle\sum_{r=1}^{\infty} A_n) = \sum_{r=1}^{\infty} \mathbb{P}(G \cap A_r)$$

$$= \textstyle\sum_{r=1}^{\infty} \mathbb{P}(G)\mathbb{P}(A_r)$$

$$= \mathbb{P}(G)\mathbb{P}(\textstyle\sum_{r=1}^{\infty} A_r).$$

Hence, $\widetilde{\mathcal{D}}_1$ indeed is Dynkin. Now, because $\mathcal{G}_1 \subseteq \widetilde{\mathcal{D}}_1$, then $\mathcal{D}(\mathcal{G}_1) \subseteq \widetilde{\mathcal{D}}_1$ and therefore, $\mathcal{D}(\mathcal{G}_1)$ as a subsystem of an independent family, is independent. Next, we consider the independent set $\{\mathcal{D}(\mathcal{G}_1), \mathcal{G}_2, \ldots, \mathcal{G}_n\}$ and do the same with $\mathcal{G}_2$ by enhancing it to $\mathcal{D}(\mathcal{G}_2)$ and so on until we exhaust the whole set. We can use induction if we want to. Thus, the statement is proved. □

2.1 Let μ^* be an outer measure on $\mathcal{P}(\Omega)$, $\mu_0^* = \mathrm{Res}_{\Sigma^*}\mu^*$, and ν^* be the outer measure induced by μ_0^*. Show that $\mu^* \leq \nu^*$ on $\mathcal{P}(\Omega)$.

Solution. If $\nu^*(Q) < \infty$, then, for every $\varepsilon > 0$, there is a cover $\{A_n\} \in \mathscr{C}_Q(\Sigma^*)$ with

$$\mu^*(Q) \leq \mu_0^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \textstyle\sum_{n=1}^{\infty} \mu_0^*(A_n) \leq \nu^*(Q) + \varepsilon.$$

This yields the desired inequality. If $\nu^*(Q) = \infty$, then the inequality $\mu^*(Q) \leq \nu^*(Q)$ holds immediately. □

2.2 Let $(\mathcal{G}, \gamma)$ be a formatter of the outer measure μ^* defined by (2.8). Show that if γ is an elementary content and $\mathcal{G}$ is a semiring, then for each $G \in \mathcal{G}$, there is a cover $\{C_n\}$ of G such that $G = \sum_{n=1}^{\infty} C_n$ and $\mu^*(G) = \sum_{n=1}^{\infty} \gamma(C_n)$.

Solution. Take $G \in \mathcal{G}$. Thus $\mathscr{C}_G \neq \emptyset$ and

$$\mu^*(G) = \inf\{\textstyle\sum_{n=1}^{\infty} \gamma(G_n) \colon \{G_n\} \in \mathscr{C}_G\}.$$

Let $\{A_n\} \in \mathscr{C}_G$. Then, by Proposition 1.3(iii) and Corollary 1.4(i) there is a countable decomposition $\sum_{n=1}^{\infty} C_n$ of G such that

$$\textstyle\sum_{n=1}^{\infty} \gamma(C_n) \leq \sum_{n=1}^{\infty} \gamma(A_n).$$

Hence,

$$\textstyle\sum_{n=1}^{\infty} \gamma(C_n) = \inf\{\sum_{n=1}^{\infty} \gamma(G_n) \colon \{G_n\} \in \mathscr{C}_G\} = \mu^*(G). \qquad \square$$

2.6 Let μ^* be an outer measure on $\mathcal{P}(\Omega)$ and $\{A_n\}$ a sequence of disjoint μ^*-measurable sets. Show that for any $Q \subseteq \Omega$,

$$\mu^*(Q \cap \textstyle\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu^*(Q \cap A_n).$$

Solution. By subadditivity,

$$\mu^*(Q \cap \textstyle\sum_{n=1}^{\infty} A_n) = \mu^*(\sum_{n=1}^{\infty} Q \cap A_n) \leq \sum_{n=1}^{\infty} \mu^*(Q \cap A_n).$$

So, we need to show the inverse inequality. Equation (2.3g) in Theorem 2.3 applied to a set $\widetilde{Q} \subseteq \Omega$ and $\{A_n\}$ reads

$$\mu^*(\widetilde{Q}) \geq \textstyle\sum_{k=1}^{\infty} \mu^*(\widetilde{Q} \cap A_k) + \mu^* \left[\widetilde{Q} \cap (\sum_{k=1}^{\infty} A_k)^c\right].$$

Replacing $\widetilde{Q}$ by $Q \cap \sum_{n=1}^{\infty} A_n$ in the last inequality we have

$$\mu^*(Q \cap \textstyle\sum_{n=1}^{\infty} A_n)$$

$$\geq \sum_{k=1}^{\infty} \mu^* \Big(Q \cap \underbrace{(\sum_{n=1}^{\infty} A_n) \cap A_k}_{= A_k} \Big)$$

$$+ \underbrace{\mu^*(Q \cap (\sum_{n=1}^{\infty} A_n) \cap (\sum_{k=1}^{\infty} A_k)^c)}_{= 0},$$

which leads to the desired inverse inequality. □

2.8 Show that $\overline{\Sigma}$ defined in Definition 2.5(ii) is a σ-algebra, $\overline{\mu}$ is a measure, that this extension does not depend upon representations of sets of $\overline{\Sigma}$, and that $(\Omega, \overline{\Sigma}, \overline{\mu})$ is complete.

Solution.

(i) We show that $\overline{\Sigma}$ is a σ-algebra. Firstly, $\Sigma \subseteq \overline{\Sigma}$. This is obvious, because if $A \in \Sigma$, then $A = A \cup \emptyset = A \in \overline{\Sigma}$. In particular, $\Omega \in \overline{\Sigma}$.

Next, let $\overline{A} \in \overline{\Sigma}$. Then, $\overline{A} = A \cup M$, where $A \in \Sigma$ and M is a subset of some μ-null set, say N. Hence,

$$\overline{A}^c = A^c \backslash M = (A^c \backslash N) + A^c \cap (N \backslash M).$$

This representation of A^c shows that it is a union of a measurable set $A^c \backslash N$ and a negligible set $A^c \cap (N \backslash M)$. Thus, $\overline{A}^c \in \overline{\Sigma}$.

Finally, Let $\{\overline{A}_n\} \subseteq \overline{\Sigma}$, where $\overline{A}_n = A_n \cup M_n$, such that $A_n \in \Sigma$ and M_n is negligible, $n = 1, 2, \ldots$. Then,

$$\bigcup_{n=1}^{\infty} \overline{A}_n = \left(\bigcup_{n=1}^{\infty} A_n \right) \cup \left(\bigcup_{n=1}^{\infty} M_n \right).$$

Obviously, the set $\bigcup_{n=1}^{\infty} M_n$ is negligible, which makes the last representation of $\bigcup_{n=1}^{\infty} \overline{A}_n$ $\overline{\Sigma}$-measurable.

(ii) Let

$$[A] : = \{A \cup M : A \in \Sigma \text{ and } M\text{'s are all negligible sets}\}.$$

Then, by the definition of $\overline{\mu}$ [see Definition 2.5(ii)],

$$\overline{\mu}_*([A]) = \mu(A)$$

which shows that $\overline{\mu}$ is invariant of a representation of a $\overline{\Sigma}$-measurable set.

(iii) We show that for any $N \in \mathcal{N}_\mu$, $\mathcal{P}(N) \subseteq \overline{\Sigma}$. Let $M \subseteq N$. From the definition of $\overline{\Sigma}$, for each $A \in \Sigma \ \Rightarrow A \cup M \in \overline{\Sigma}$. We can take the empty set as A to see that $M \in \overline{\Sigma}$. □

2.10 Let μ^* be an outer measure on $\mathcal{P}(\Omega)$ and $N \subseteq \Omega$ be such that $\mu^*(N) = 0$. Show that for any subset $Q \subseteq \Omega$, $\mu^*(Q \cup N) = \mu^*(Q)$.

Solution. Obviously,

$$\mu^*(Q \cup N) \leq \mu^*(Q) + \mu^*(N) = \mu^*(Q).$$

The inverse inequality follows from the monotonicity of the outer measure applied to $Q \cup N$ and its subset Q. □

2.12 Let μ^* be the outer measure generated by an extendible σ-finite formatter $(\mathcal{G}, \gamma)$, Σ^* be the σ-algebra of all μ^*-measurable sets, and let $\Sigma(\mathcal{G})$ be the σ-algebra generated by $\mathcal{G}$. Show that for every $A^* \in \Sigma^*$, there is a set $B \in \Sigma(\mathcal{G})$ with $B \subseteq A^*$ and $\mu^*(A^* \backslash B) = 0$.

Solution. Because $(A^*)^c \in \Sigma^*$ and $(\mathcal{G}, \gamma)$ is σ-finite, by Corollary 2.18, there is a set $B^c \in \Sigma$ such that $B^c \supseteq (A^*)^c$ and $\mu^*(B^c \backslash (A^*)^c) = 0$. The assertion of the problem follows when noticing that $B^c \backslash (A^*)^c = A^* \backslash B$. □

2.14 Let Σ be a σ-algebra in Ω and let $a \in \Omega$. Show that for $\{a\} \in \Sigma$, the measure space $(\Omega, \Sigma, \varepsilon_a)$ is complete if and only if $\Sigma = \mathcal{P}(\Omega)$.

Solution. (i) If $\Sigma = \mathcal{P}(\Omega)$, then $(\Omega, \Sigma, \varepsilon_a)$ is obviously complete.

(ii) Suppose Σ is a proper subset of $\mathcal{P}(\Omega)$. We show that $(\Omega, \Sigma, \varepsilon_a)$ is incomplete. Denote $\Omega' = \{a\}^c$ which is an element of Σ and it is a null set of ε_a. Then, $\varepsilon_a(\Omega') = 0$. Now, we show that there is a nonmeasurable subset of Ω'. Suppose it is not true and that all subsets of Ω' are measurable. Note that

$$\mathcal{P}(\Omega) = \mathcal{P}(\Omega') \cup \{A \subseteq \Omega : A = A' \cup \{a\} : A' \in \mathcal{P}(\Omega')\}.$$

Because by the assumption, Σ is a proper subset of $\mathcal{P}(\Omega)$, there is a nonmeasurable subset of Ω, say P. Then, either $P \subseteq \Omega'$ or $P = Q' \cup \{a\}$ where $Q' \subseteq \Omega'$. Clearly, $Q' \notin \Sigma$, because otherwise P would. Thus, in either case, Ω' has a nonmeasurable subset. □

2.16 Let (Ω, Σ, μ) be a measure space, $\mathcal{G}_\infty = \{G \in \Sigma : \mu(G) < \infty\}$, and $\Sigma_\infty = \{Q \subseteq \Omega : Q \cap G \in \Sigma, \forall G \in \mathcal{G}_\infty\}$. Define the set function μ_∞ on Σ_∞ as

$$\mu_\infty(Q) = \begin{cases} \mu(Q), & Q \in \Sigma \\ \infty, & Q \notin \Sigma. \end{cases}$$

Problem 1.13 claimed that μ_∞ was a measure on Σ_∞. Show that if the measure space (Ω, Σ, μ) is complete, then so is μ_∞.

Solution. Let N_∞ be a μ_∞-negligible set. Then, there is a Σ_∞-measurable μ_∞-null superset N of N_∞. This N must also belong to Σ [because if it would not, $\mu_\infty(N)$ would be equal to ∞], on which μ and μ_∞ coincide. Consequently, N_∞ is also a μ-negligible set and because (Ω, Σ, μ) is complete, N_∞ is in Σ and thus in Σ_∞. □

2.18 Let μ^* be the outer measure generated by an extendible formatter $(\mathcal{G}, \gamma)$ on a nonempty set Ω, Σ^* be the σ-algebra of all μ^*-measurable sets, and $\Sigma(\mathcal{G})$ be the σ-algebra generated by $\mathcal{G}$. Show that a subset $N \subseteq \Omega$ is negligible if and only if $\mu^*(N) = 0$.

Solution. Denote $\mu_0^* = \text{Res}_{\Sigma^*}\mu^*$ and $\mu = \text{Res}_{\Sigma(\mathcal{G})}\mu_0^*$.

(a) Suppose N is negligible. Then, there is a μ-null superset of N, say N_0. Consequently, $\mu^*(N_0) = 0$. By monotonicity of μ^*,

$$\mu^*(N) = 0.$$

(b) Let $N \subseteq \Omega$ be such that $\mu^*(N) = 0$. First, we show that $N \in \Sigma^*$. Let $Q \subseteq \Omega$. Then, by (2.3) and due to $\mu^*(Q \cap N) = 0$,

$$\mu^*(Q) \leq \mu^*(Q \cap N) + \mu^*(Q \cap N^c) = \mu^*(Q \cap N^c) \leq \mu^*(Q),$$

which yields

$$\mu^*(Q) = \mu^*(Q \cap N^c)$$

and hence,

$$\mu^*(Q) = \mu^*(Q \cap N) + \mu^*(Q \cap N^c).$$

That $N \in \Sigma^*$ yields $\mu_0^*(N) = 0$ and, in particular, $\mu_0^*(N) < \infty$. By Lemma 2.17, there is an $N_0 \supseteq N$ such that $N_0 \in \Sigma(\mathcal{G})$ and

$$\mu_0^*(N_0 \backslash N) = 0.$$

The latter leads to

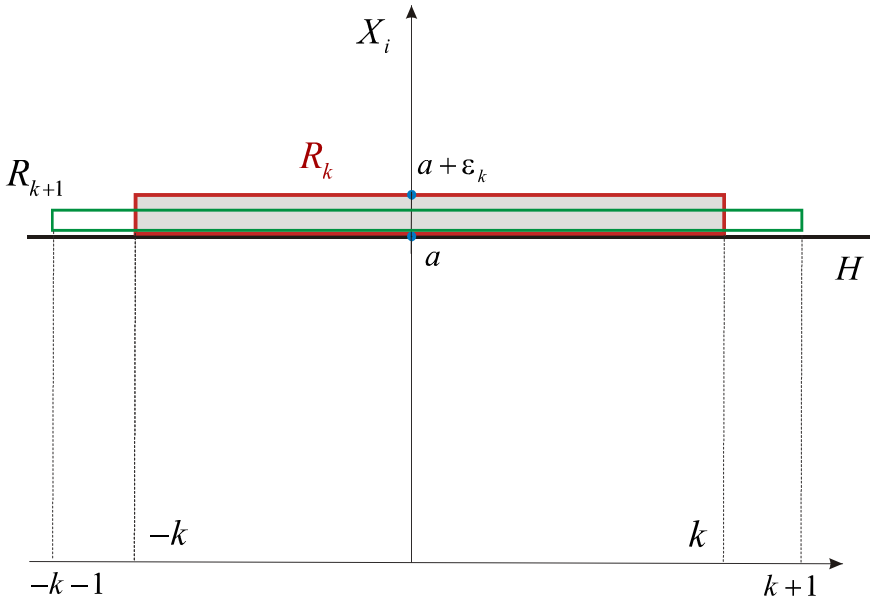
$$\mu(N_0) = \mu_0^*(N_0) = \mu_0^*(N) = 0$$

and thus implies that N is negligible. □

3.1 Let $H = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n\colon\ x_i = a \in \mathbb{R}\}$ be a hyperplane orthogonal to the ith coordinate axis. Show that H is a λ-null Borel set. [*Hint*: 1) Show that H is closed in $(\mathbb{R}^n, \tau_e)$ and hence Borel; 2) find a relevant countable cover of H by rectangles from $\mathcal{S}$ and apply Lemma 3.6.]

Solution. Clearly, H^c is open. Indeed, let $x \in \mathbb{R}^n$ such that the ith coordinate $x_i = b > a$. Then, the distance from x to H is exactly $b - a$. Thus, if $B(x, r)$ is an open ball with radius $r < b - a$, $B(x, r) \subseteq H^c$. The same holds true with a point $y \in \mathbb{R}^n$ whose ith coordinate $y_i = b < a$. Thus H is closed and consequently, H is

Borel. Let $R_k := (\alpha_k, \beta_k] \in \mathbb{R}^n$ be a half-open rectangle located upon H such that all but its ith coordinate are $-k$ for α_k and k for β_k. The ith left coordinate of R_k is a and the right coordinate of R_k is $a + \varepsilon_k$ such that $\varepsilon_k = 2^{-k}(2k)^{1-n}\varepsilon$, $\varepsilon > 0$.



Obviously, $H \subseteq R := \bigcup_{k=1}^{\infty} R_k$. We now calculate $\lambda^n(R)$:

$$\lambda^n(R_k) = \left(\prod_{k=1}^{n-1}(k-(-k))\right) \times \left(a + 2^{-k}(2k)^{1-n}\varepsilon - a\right) = 2^{-k}\varepsilon$$

$$\Rightarrow$$

$$\lambda^n(H) \le \lambda^n(R) \le \textstyle\sum_{k=1}^{\infty}\lambda^n(R_k) = \varepsilon.$$

Thus, by Lemma 3.6,

3.6 Lemma. *A set $N \subseteq \mathbb{R}^n$ is λ-negligible if and only if for each $\varepsilon > 0$ there is a countable cover by half-open intervals $\{I_k\} \subseteq \mathcal{S}$ of N such that $\sum_{k=1}^{\infty}\lambda^0(I_k) < \varepsilon$.*

$\lambda^n(H) = 0.$ □

3.3 Show that f defined by (3.5) in Remark 3.5(iii) is an extended distribution function.

Solution. We assume that $R_n \subseteq C,\ n = 1, 2, \ldots,$ where C is a compact set in $(\mathbb{R}, \tau_e)$. Recall that any "figure" $R_n \in \mathcal{R}$ is a finite union of disjoint intervals of type $(a, b]$. Because of the right-continuity of f, it can be easily shown that, for each fixed $\varepsilon > 0$ for any figure $R_n,$ there is a subfigure $B_n \subseteq R_n$ such that $\overline{B}_n \subseteq R_n$ and such that

$$\mu_f(R_n) - \mu_f(B_n) < \varepsilon 2^{-n}.$$

It also follows that $\bigcap_{n=1}^{\infty} \overline{B}_n = \emptyset$. We claim that there is an r such that the intersection $\bigcap_{k=1}^{r} \overline{B}_k = \emptyset$. To see this, observe that

$$\{C \backslash \overline{B}_n = C \cap (\overline{B}_n)^c;\ N = 1, 2, \ldots\}$$

is an open cover of C in the relative topology $(C,\ \tau_e \cap C)$. Because compactness is weakly hereditary and C is closed, it follows that C is also compact in $\tau_e \cap C$. Thus, the above cover reduces to a finite subcover, for example, $C \backslash \overline{B}_1, \ldots, C \backslash \overline{B}_r$ yielding

$$C = \bigcup_{k=1}^{r} (C \backslash \overline{B}_k) = C \backslash \left(\bigcap_{k=1}^{r} \overline{B}_k \right).$$

Thus,

$$\bigcap_{k=1}^{r} \overline{B}_k = \emptyset \quad \text{and} \quad \bigcap_{k=1}^{r} B_k = \emptyset.$$

Now, for all $n \geq r$,

$$\bigcap_{k=1}^{n} B_k = \emptyset$$

and

$$\mu_f(R_n) = \mu_f(R_n \setminus \bigcap_{k=1}^{n} B_k) + \mu_f(\bigcap_{k=1}^{n} B_k)$$

$$= \mu_f\left(R_n \setminus \bigcap_{k=1}^{n} B_k\right).$$

In as much as $\{R_n\}$ is monotone decreasing it follows that

$$R_n \setminus \left(\bigcap_{k=1}^{n} B_k\right) = \bigcup_{k=1}^{n} (R_n \setminus B_k) \subseteq \bigcup_{k=1}^{n} (R_k \setminus B_k).$$

Observe that this is the desired inclusion implying the estimate $\mu_f(R_n) < \varepsilon$. The above inclusion was due to the inclusion $R_n \setminus B_k \subseteq R_k \setminus B_k$ which holds for all $k \leq n$ (as long as $n < \infty$). Hence the above countable intersection reduces to a finite intersection of the sets $\overline{B}_k,\ k = 1, \ldots, r$. Thus we have

$$\mu_f(R_n) = \mu_f\left(R_n \setminus \left(\bigcap_{k=1}^{n} B_k\right)\right)$$

$$\leq \mu_f\left(\bigcup_{k=1}^{n} R_k \setminus B_k\right) \leq \sum_{k=1}^{n} \mu_f(R_k \setminus B_k)$$

$$\leq \varepsilon(1 - 2^{-n})$$

which shows that $\mu_f(R_n) \to 0$. □

3.4 Let f_1 and f_2 be two extended distribution functions and let μ_1 and μ_2 be the corresponding Borel - Lebesgue - Stieltjes measures induced by these functions. Show that $\mu_1 = \mu_2$ if and only if $f_1 - f_2 = \boldsymbol{c},$ where $\boldsymbol{c}$ is a constant function.

Solution. (i) Suppose f_1 and f_2 differ in a constant. Define $\mu_{f_i}((a, b]) = f_i(b) - f_i(a)$. Therefore $\mu_{f_1} = \mu_{f_2}$ on semi-ring $\mathcal{S}$ and thus obviously on the ring $\mathcal{R}(\mathcal{S})$. As Lebesgue - Stieltjes contents, they have all conditions met for a unique Carathéodory extension of the corresponding formatters from the ring $\mathcal{R}(\mathcal{S})$ in $\mathbb{R}$ to the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. See also Remark 3.5$(i$-$iii)$.

(ii) Conversely, if μ_1 and μ_2 are two Borel - Lebesgue - Stieltjes measures induced by two extended distribution functions f_1 and f_2, such that μ_1 and μ_2 coincide on semi-ring $\mathcal{S}$, then we have for a fixed a and any $x > a$,

$$\mu_1((a, x]) = f_1(x) - f_2(a)$$

$$\mu_2((a, x]) = f_2(x) - f_2(a).$$

Therefore, $f_1(x) = f_2(x) + c$, where $c = f_1(a) - f_2(a)$.

For $x < a$,

$$\mu_1((x, a]) = f_1(a) - f_1(x)$$

and

$$\mu_2((x, a]) = f_2(a) - f_2(x).$$

So,

$$-f_1(a) + f_1(x) + f_2(a) - f_2(x) = 0$$

and

$$f_1(x) = f_2(x) + c,$$

where $c = f_1(a) - f_2(a)$, the same number. □

3.9 Let f be a distribution function and let μ_f denote the Borel - Lebesgue - Stieltjes measure induced by f. Justify with a proof or give a counter-argument:

a) Must f be an extended distribution function?

b) Suppose g is a function defined by $(*)$ of Remark 3.5 (iii). Is g a distribution function? If your answer is "yes," is $g = f$?

Solution. In Remark 3.5(iii), definition (3.5) of an extended distribution function may look confusing, primarily because we set

$f(0) = 0$. However, we recall that (3.5) defines only one representative of the entire class of distribution functions. The rest of them can be restored by adding to f the set $\mathbb{R}$ of all real constants. Thus, $\mathfrak{D}_b \subseteq \mathfrak{D}$, even though $f(0)$ need not be equal to 0. Furthermore, if we employ (3.5) by means of μ_f, then the obtained function this way is an extended distribution function, say h, which is equal to zero at the origin, but it need not be a distribution function, as it need not vanish at $-\infty$. However, f and h differ in a constant.

a) Therefore, the answer to a) is "yes."

b) As pointed out, g need not be a distribution function. However, if it is one, it must necessarily equal f. □

3.12 Let $(\Omega, \mathcal{B}, \mu)$ be a Borel measure space, such that the Borel σ-algebra $\mathcal{B}$ is generated by a Hausdorff topological space τ, and μ is a finite Borel measure on $\mathcal{B}$. For any subset $Q \subseteq \Omega$, denote by $\mathcal{K}(Q)$ the collection of all compact subsets of Q. Show that the family $\mathcal{M} \subseteq \mathcal{B}$ of all Borel sets B such that

$$\mu(B) = \sup\{\mu(K)\colon\ K \in \mathcal{K}(B)\}$$

is a monotone system in Ω, that is, all Borel sets that can be approximated from below by compact Borel subsets form a monotone system.

Solution. We need to show that for each monotone increasing sequence $\{M_n\}\uparrow\ \subseteq \mathcal{M}$, $\bigcup\limits_{n=1}^{\infty} M_n \in \mathcal{M}$, and for each monotone decreasing sequence $\{M_n\}\downarrow\ \subseteq \mathcal{M}$, $\bigcap\limits_{n=1}^{\infty} M_n \in \mathcal{M}$.

1) Let $\{M_n\}\uparrow\ \subseteq \mathcal{M}$. Then $\forall\ \varepsilon/2 > 0$ and $\forall n = 1, 2, \ldots,$ there is $K_n(\varepsilon) \in \mathcal{K}(M_n)$:

$$\mu(M_n) \leq \mu(K_n(\varepsilon)) + \varepsilon/2.$$

Let $C_n(\varepsilon) := \bigcup_{k=1}^{n} K_k(\varepsilon)$, then $C_n(\varepsilon)$ is also a compact set with $K_n(\varepsilon) \subseteq C_n(\varepsilon) \subseteq M_n$, which yields

$$\mu(M_n) \leq \mu(C_n(\varepsilon)) + \varepsilon/2.$$

By continuity from below (Lemma 1.6) we have

$$\mu(M) = \lim_{n \to \infty} \mu(M_n)$$

$$\leq \lim_{n \to \infty} \mu(C_n(\varepsilon)) + \varepsilon/2 = \mu(C(\varepsilon)) + \varepsilon/2,$$

where $C = \bigcup_{n=1}^{\infty} C_n$. Because $\{C_n(\varepsilon)\}$ is a monotone increasing sequence of compacts subsets of M, for any $\varepsilon/2$ there is an $N \in \mathbb{N}$ such that, $\forall n \geq N,\ \mu(C(\varepsilon)) \leq \mu(C_n(\varepsilon)) + \varepsilon/2$, what finally yields

$$\mu(M) \leq \mu(C_n(\varepsilon)) + \varepsilon.$$

In other words, $\forall \varepsilon > 0$ there is a compact subset of M satisfying the above inequality. This is equivalent to

$$\mu(M) = \sup\{\mu(K) : K \in \mathcal{K}(M)\}$$

$\Rightarrow\ M \in \mathcal{M}$.

2) Let $\{M_n\}\downarrow\ \subseteq \mathcal{M}$. Then $\forall \varepsilon > 0\ \forall n = 1, 2, \ldots,$

$$\exists\, K_n(\varepsilon) \in\ \mathcal{K}(M_n) : \mu(M_n) \leq \mu(K_n(\varepsilon)) + \tfrac{\varepsilon}{2^n}\,.$$

Let $K(\varepsilon) :\ = \bigcap_{n=1}^{\infty} K_n(\varepsilon)$. According to Problem 6.10, Chapter 3, any intersection of compact subsets in a Hausdorff topological space is compact. Thus, $K(\varepsilon)$ is also compact and in addition, $K(\varepsilon) \subseteq M$. Because μ is finite, $\mu(M \backslash K(\varepsilon)) = \mu(M) - \mu(K(\varepsilon))$. On the other hand,

$$M\backslash K(\varepsilon) = \left[\bigcap_{n=1}^{\infty} M_n\right] \backslash \left[\bigcap_{i=1}^{\infty} K_i(\varepsilon)\right]$$

$$= \left[\bigcap_{n=1}^{\infty} M_n\right] \cap \left[\bigcup_{i=1}^{\infty} K_i^c(\varepsilon)\right]$$

$$= \bigcap_{n=1}^{\infty} \left(M_n \cap \left[\bigcup_{i=1}^{\infty} K_i^c(\varepsilon)\right]\right)$$

$$= \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{\infty} (M_n \cap K_i^c(\varepsilon))$$

$$= \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{\infty} (M_n \backslash K_i(\varepsilon)). \tag{P3.12}$$

Now, if $\{A_i : i \in I\}$ and $\{B_j : j \in J\}$ are two indexed families of sets, we obviously have:

$$\bigcup_{i\in I} \bigcap_{j\in J} (A_i \cap B_j) = \bigcup_{i\in I} \left[\bigcap_{j\in J} B_j\right] \cap A_i$$

$$= \left(\bigcup_{i\in I} A_i\right) \cap \left[\bigcap_{j\in J} B_j\right] = \bigcap_{j\in J} \left[B_j \cap \left(\bigcap_{i\in I} A_i\right)\right]$$

$$= \bigcap_{j\in J} \bigcup_{i\in I} (A_i \cap B_j). \tag{P3.12a}$$

Consequently, due to (P3.12a), equation (P3.12) yields

$$M\backslash K(\varepsilon) = \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{\infty} (M_n \cap K_i^c(\varepsilon)) = \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{\infty} (M_n \backslash K_i(\varepsilon))$$

$$= \bigcup_{i=1}^{\infty} \bigcap_{n=1}^{\infty} (M_n \backslash K_i(\varepsilon)). \tag{P3.12b}$$

We show that

$$M\backslash K(\varepsilon) \subseteq \bigcup_{k=1}^{\infty} (M_k\backslash K_k(\varepsilon)). \tag{P3.12c}$$

Indeed, if $M\backslash K(\varepsilon) \neq \emptyset$ (which otherwise would be trivial), then, due to (P3.12b), for any $\omega \in M\backslash K(\varepsilon)$, there is an i_0 such that $\omega \in \bigcap_{n=1}^{\infty} (M_n\backslash K_{i_0}(\varepsilon))$. Therefore, $\omega \in M_{i_0}\backslash K_{i_0}$, which in turn is a subset of $\bigcup_{k=1}^{\infty} (M_k\backslash K_k(\varepsilon))$ herewith validating inclusion (P3.12c). Consequently, we have

$$\begin{aligned}\mu(M) - \mu(K(\varepsilon)) &= \mu(M\backslash K(\varepsilon)) \\ &\leq \mu\left[\bigcup_{k=1}^{\infty} (M_k\backslash K_k(\varepsilon))\right] \\ &\leq \sum_{k=1}^{\infty} \mu(M_k) - \mu(K_k(\varepsilon)) \leq \varepsilon\end{aligned}$$

and hence, $M \in \mathcal{M}$.

So we proved that $\mathcal{M}$ is indeed a monotone system. □

3.13 Let $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ be a finite Borel measure space (that is, μ is finite). Show that μ is weakly regular, that is, it is outer and closed regular on $\mathcal{B}(\mathbb{R}^n)$.

Solution. Let $\mathcal{W} \subseteq \mathcal{B}$ be the family of all weakly regular Borel sets. That is [see Definition 3.12 (iv)] for $A \in \mathcal{W}$,

$$\mu(A) = \inf\{\mu(O) : A \subseteq O \in \tau_e\} \tag{P3.13}$$

and

$$\mu(A) = \sup\{\mu(F) : A \supseteq F \in \tau_e^c\}. \tag{P3.13a}$$

Weak regularity seems to relate to some, but surely not to all, Borel sets. However, we prove that under the above assumptions, $\mathcal{W} = \mathcal{B}(\mathbb{R}^n)$. It can be easily shown that (P3.13 - P3.13a) is equivalent to

the following assertion. Given an $\varepsilon > 0$, there is an open set O and a closed set F such that

$$F \subseteq A \subseteq O \text{ and } \mu(O \backslash F) < \varepsilon. \tag{P3.13b}$$

(i) Obviously, from the definition, open sets are outer regular.

(ii) Let F_σ be a countable union of closed sets. By Problem 1.26, Chapter 3, each open set in τ_e is an F_σ set. That is,

$$\tau_e \ni O = \bigcup_{k=1}^{\infty} F_k.$$

By Problem 1.18, Chapter 3, we can even assume that $\{F_k\}$ is a monotone nondecreasing sequence from τ_e^c (which is obvious anyway). Then, by continuity from below,

$$\mu(O) = \lim_{k\to\infty} \mu(F_k) = \sup\{\mu(F) : \tau_e^c \ni F \subseteq O\}$$

that implies that $\tau_e \subseteq \mathcal{W}$, i.e. open sets can be approximated by open sets and by closed sets (a weaker form of the closed regularity).

(iii) We show that $\mathcal{W}$ is a σ-algebra.

$a)$ $\mathbb{R}^n \in \mathcal{W}$ because it is open.

$b)$ Let $Q \in \mathcal{W}$ and an $\varepsilon > 0$ be chosen. Then, according to (P3.13b), there are open and closed sets, say O and F, respectively, with

$$F \subseteq Q \subseteq O \text{ and } \mu(O \backslash F) < \varepsilon.$$

Hence,

$$\tau_e^c \ni O^c \subseteq Q^c \subseteq F^c \in \tau_e$$

and

$$\mu(F^c \backslash O^c) = \mu(\mathbb{R}^n) - \mu(F) - \mu(\mathbb{R}^n) + \mu(O) < \varepsilon$$

implying that $Q^c \in \mathcal{W}$.

c) Let $\{Q_k\} \subseteq \mathcal{W}$ and $\varepsilon > 0$ be chosen. Then, for each k, there is a couple of sets, $F_k \in \tau_e^c$ and $O_k \in \tau_e$, such that

$$F_k \subseteq Q_k \subseteq O_k \text{ and } \mu(O_k \backslash F_k) < \varepsilon/2^k.$$

Let $O = \bigcup_{k=1}^{\infty} O_k$ and $F_\sigma = \bigcup_{k=1}^{\infty} F_k$. Then, because of

$$O \backslash F_\sigma = \bigcup_{k=1}^{\infty} \left[O_k \cap \left(\bigcap_{j=1}^{\infty} F_j^c \right) \right]$$

$$\subseteq \bigcup_{k=1}^{\infty} [O_k \cap F_k^c] = \bigcup_{k=1}^{\infty} (O_k \backslash F_k),$$

we get

$$\mu(O \backslash F_\sigma) \leq \mu\left(\bigcup_{k=1}^{\infty} O_k \backslash F_k \right) \leq \textstyle\sum_{k=1}^{\infty} \mu(O_k \backslash F_k) < \varepsilon. \quad \text{(P3.13c)}$$

Although F_σ is not closed, it is an F_σ set itself, and as such it can be represented as the countable union of a monotone nondecreasing sequence, say $\{C_k\}$, of closed sets. In fact, this sequence can be chosen as $\left\{ \bigcup_{j=1}^{k} F_j \right\}$. By continuity from below we then have

$$\mu(F_\sigma) = \lim_{k \to \infty} \mu\left(\bigcup_{j=1}^{k} F_j \right)$$

implying that for the same $\varepsilon > 0$ chosen earlier, there is an N such that $\forall n \geq N$,

$$\mu(F_\sigma) - \mu\left(\bigcup_{j=1}^{k} F_j \right) < \varepsilon. \quad \text{(P3.13d)}$$

Therefore, from (P3.13c) and (P3.13d) we have

$$\mu\left(O \backslash \bigcup_{j=1}^{k} F_j \right) = \mu(O) - \mu\left(\bigcup_{j=1}^{k} F_j \right)$$

$$= \mu(O) - \mu(F_\sigma) + \mu(F_\sigma) - \mu\left(\bigcup_{j=1}^{k} F_j\right) \leq 2\varepsilon,$$

with

$$\tau_e \ni \bigcup_{j=1}^{k} F_j \subseteq \bigcup_{j=1}^{\infty} F_j = F_\sigma \subseteq \bigcup_{k=1}^{\infty} Q_k \subseteq \bigcup_{k=1}^{\infty} O_k \in \tau_e,$$

concluding that $\bigcup_{k=1}^{\infty} Q_k \in \mathcal{W}$ and that $\mathcal{W}$ is indeed a σ-algebra.

(iv) Because $\mathcal{W}$ is a σ-algebra and $\tau_e \subseteq \mathcal{W}$, it follows that $\mathcal{B} \subseteq \mathcal{W}$ and thus $\mathcal{W} = \mathcal{B}$. □

3.18 Show that a σ-finite Borel measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ need not be outer regular.

Solution. Let μ be defined as follows. Let $B := \left\{1, \frac{1}{2}, \dots\right\}$ and $\forall$ $n = 1, 2, \dots,\ \mu(\{\frac{1}{n}\}) = \frac{1}{n}$, while for every subset A of B^c we let $\mu(A) = 0$. If $B_n := \left\{1, \frac{1}{2}, \dots, \frac{1}{n}\right\}$ then obviously $\mu(B_n \cup B^c)$ is finite and $B_n \cup B^c \uparrow \mathbb{R} = \Omega$. Although $\mu(\mathbb{R}) = \mu(B) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$, μ is finite on $B_n \cup B^c \Rightarrow \mu$ is σ-finite. We show that Borel sets cannot be μ-approximated by open supersets in general. Let us take set $\{0\}$. If O is any open set containing 0, then O must also contain an open interval centered at zero and thus, for some $\varepsilon > 0$,

$$\mu(O) \geq \mu((-\varepsilon, +\varepsilon)) \geq \mu\left(\sum_{n=k}^{\infty} \frac{1}{n}\right) = \infty \text{ (for some } k).$$

Thus, $\mu(O) = \infty$ holding for any open set containing set $\{0\}$, whereas $\mu(\{0\}) = 0$. □

3.19 Let $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$ be the Borel - Lebesgue measure space. Show that the Borel - Lebesgue measure λ is strongly regular.

Solution. From Definition 3.12$(iv$-$v)$, we need to show that λ is

(A) finite on $\mathcal{K}(\mathbb{R}^n)$ (which is obvious, because in $(\mathbb{R}^n, \tau_e)$, compact sets are d_e-bounded and λ is a Borel - Lebesgue - Stieltjes measure)

(B) inner regular (it is done as per Problem 3.15, because λ is a σ-finite Borel measure)

(C) outer regular, i.e., for each Borel set $A \in \mathcal{B}(\mathbb{R}^n)$,

$$\lambda(A) = \inf\{\mu(O) : A \subseteq O \in \tau_e\}. \tag{P3.19}$$

The latter is what we are going to work on in this problem. Firstly assume that, given an $A \in \mathcal{B}, \lambda(A) < \infty$, because otherwise, the result is trivial considering monotonicity of λ.

We need to show that for each $\varepsilon > 0$, there is an open set $G \in \tau_e$ such that

$$\lambda(A) + \varepsilon > \lambda(G). \tag{P3.19a}$$

By Carathéodory's extension, A can be approximated by half-open rectangles. More precisely, if $\mathscr{C}_A = \mathscr{C}_A(\mathcal{S}^n)$ is a the set of all at most countable covers of A by half-open rectangles (i.e., elements of the semi-ring $\mathcal{S}^n$), then

$$\lambda(A) = \inf\{\textstyle\sum_{r=1}^{\infty}\lambda(R_r) : \{R_r\} \in \mathscr{C}_A\}.$$

Therefore, for each $\varepsilon > 0$, there is a cover, say $\{R_k\} \in \mathscr{C}_A$, such that

$$\lambda(A) + \varepsilon\tfrac{1}{2} > \textstyle\sum_{k=1}^{\infty}\lambda(R_k). \tag{P3.19b}$$

In turn, given ε, for each R_k there is an open superset, in fact a rectangle, say G_k, such that

$$\lambda(R_k) + \varepsilon/2^{k-1} > \lambda(G_k). \tag{P3.19c}$$

If $G = \bigcup\limits_{k=1}^{\infty} G_k$, then G is an open superset of A. It remains to show that G has the required property (P3.19a). Indeed, from (P3.19b-c)

$$\lambda(A) + \tfrac{\varepsilon}{2} > \textstyle\sum_{k=1}^{\infty}\lambda(G_k) - \tfrac{\varepsilon}{2} \geq \lambda(G) - \tfrac{\varepsilon}{2}$$

which holds due to σ-subadditivity of λ. □

3.21 Show that if N is a negligible set in $(\mathbb{R}^n, \mathcal{B}, \lambda)$, for each $\varepsilon > 0$, there is a countable cover of N by disjoint half-open cubes $\{C_k\}$ such that

$$\sum_{k=1}^{\infty} \lambda^0(C_k) < \varepsilon.$$

Solution. By Definition 2.5(i), if N is negligible, there is a Borel λ-null superset, say N_0, of N, that is, $N_0 \in \mathcal{B}(\mathbb{R}^n)$ such that $N_0 \supseteq N$ and $\lambda(N_0) = 0$. By Problem 3.20, given an $\varepsilon > 0$, there is a countable cover of N_0 by disjoint half-open cubes $\{C_k\}$, such that

$$\sum_{k=1}^{\infty} \lambda_e(C_k) < \varepsilon.$$ □

3.24 Let μ be a translation-invariant Borel measure on $\mathcal{B}(\mathbb{R}^n)$ and let μ^* be the outer measure produced by $(\mathcal{B}(\mathbb{R}^n), \mu)$ and $\mathcal{B}^*_\mu$ be the corresponding σ-algebra of μ^*-measurable sets. Show that

(i) $\mu^* = \mu(C)\lambda^*$ on $\mathcal{P}(\mathbb{R}^n)$ and $\mu^*_0 = \mu(C)\lambda^*_0$ on $\mathcal{B}^*_\mu$, where C is a unit cube.

(ii) $\mathcal{B}^*_\mu = \mathcal{L}^*$.

Solution.

(i) By Theorem 3.10, $\mu = \mu(C)\lambda$ on $\mathcal{B}(\mathbb{R}^n)$. Let μ^*_1 be the outer measure produced by $\mu(C)\lambda$ and $\mu^*_2 = \mu(C)\lambda^*$. Because μ^*, μ^*_1, and μ^*_2 coincide on $\mathcal{B}(\mathbb{R}^n)$, by Remark 3.3($i$), $\mu^* = \mu^*_1 = \mu^*_2$ on $\mathcal{P}(\mathbb{R}^n)$. In particular,

$$\mu^*_0 = \mu(C)\lambda^*_0 \text{ on } \mathcal{B}^*_\mu.$$

(ii) In light of (i), we take $\mu(C)\lambda^*$ as a version of μ^*. Let $Q \in \mathcal{P}(\mathbb{R}^n)$. Now, the equation

$$\mu^*(Q) = \mu^*(Q \cap B) + \mu^*(Q \cap B^c)$$

holds true if and only if

$$\lambda^*(Q) = \lambda^*(Q \cap B) + \lambda^*(Q \cap B^c)$$

does. Hence $B \in \mathcal{B}_\mu^*$ if and only if $B \in \mathcal{L}^*$. □

3.28 Show that there is a non-Borel set in $\mathcal{P}(\mathbb{R})$.

Hints: We call $x, y \in \mathbb{R}$ equivalent $(x \sim y)$ if and only if $x - y \in \mathbb{Q}$ (rational numbers). For every real number $x \in \mathbb{R}$ we assign another real number y to the class A_x if and only if $x - y \in \mathbb{Q}$.

1) Show that $(\mathbb{R}, \sim)$ is indeed an equivalence relation.

Let $\mathcal{E}|_\sim$ be the quotient set of modulo $\sim$. Using the axiom of choice we select any element from each class of $\mathcal{E}|_\sim$ that belongs to set $(0, 1]$. Denote by A the collection of all such elements.

2) Show that such a selection is possible taking into account the axiom of choice; that is, it can be shown that $\forall x \in \mathbb{R}$, $A_x \cap (0, 1] \neq \emptyset$.

3) Show that set A has the following properties.

$$\forall q \neq r \in \mathbb{Q}, \ \{q + A\} \cap \{r + A\} = \emptyset. \tag{P3.28}$$

$\mathbb{R}$ can be restored from A as

$$\mathbb{R} = \bigcup_{q \in \mathbb{Q}} (q + A) = \sum_{q \in \mathbb{Q}} (q + A). \tag{P3.28a}$$

4) Finally, let $\widetilde{\mathbb{Q}} = \mathbb{Q} \cap (0, 1]$. Then $\bigcup_{q \in \widetilde{\mathbb{Q}}} (q + A) \subseteq (0, 2]$. If A is a Borel set (and this is the assumption that will lead to a contradiction), then by Problem 3.11, $x + A$ is a Borel set too; and by the translation-invariance of Borel - Lebesgue measure λ, $\lambda(x + A) = \lambda(A)$ implying that

$$\lambda\Big(\bigcup_{q\in\widetilde{\mathbb{Q}}}(q+A)\Big)=\sum_{q\in\widetilde{\mathbb{Q}}}\lambda(q+A)\leq\lambda((0,2])=2.$$

Thus the above series is finite; and because the $\lambda(q+A)$ values are equal for all $q\in\widetilde{\mathbb{Q}}$, each of them must be zero, which implies that $\lambda(q+A)=0,\ \forall q\in\mathbb{Q}$. But

$$\mathbb{R}=\sum_{q\in\mathbb{Q}}(q+A)\ \Rightarrow\ \lambda(\mathbb{R})=\sum_{q\in\mathbb{Q}}\lambda(q+A)=0,$$

which is an absurdity. Thus, our assumption that A is a Borel set was wrong.

4.3 Let (Ω,Σ,μ) be a measure space with $\Omega=\mathbb{R},\ \Sigma=\{A\subseteq\mathbb{R}:$ either A or $A^c\preceq\mathbb{N}\}$ and let $\mu(A)=0$ for $A\preceq\mathbb{N}$ and $\mu(A)=1$ for $A^c\preceq\mathbb{N}$. Let $\Omega'=\mathbb{R},\ \ \Sigma'=\mathcal{B}(\mathbb{R})$. Define $[\Omega,\Omega',f]$ as

$$f(x)=\mathbf{1}_{\mathbb{Q}^c}(x)=\begin{cases}0, & \text{if } x \text{ is rational}\\ 1, & \text{if } x \text{ is irrational.}\end{cases}$$

Prove that f is Σ-Σ'-measurable and determine μf^*.

Solution. (i) Let B be any Borel subset of $\mathbb{R}$. Then, $f^*(B)\in\{\mathbb{R},\emptyset,\mathbb{Q},\mathbb{Q}^c\}$. The latter is a σ-algebra which clearly is a subset of Σ. Thus, f is measurable.

$(ii)\ \ \mu f^*(\emptyset)=0,\mu f^*(\{0\})=\mu(\mathbb{Q})=0$, because $\mathbb{Q}$ is countable. $\mu f^*(\{1\})=\mu(\mathbb{Q}^c)=1$, because $\mathbb{Q}^c$ is uncountable. Then,

$$\mu f^*(\{0,1\})=0+1=1.$$

Now, let B be any Borel subset of $\mathbb{R}$. From (i) we have that $f^*(B)\in\{\mathbb{R},\emptyset,\mathbb{Q},\mathbb{Q}^c\}$. Hence, $\mu f^*(B)\in\{0,1\}$. □

4.5 Let $\mathcal{M}(\mathbb{R}^n)$ be the collection of all motions on $(\mathbb{R}^n,d_e)$. Show that $(\mathcal{M}(\mathbb{R}^n),\circ)$, where $\circ$ is the composition operator, forms a group.

Solution. We define a motion as $M = T_a \circ R$, where R is an orthogonal matrix and $T_a(x) = x + a$, $x, a \in \mathbb{R}^n$. [To keep consistency in notation we will write $A(x)$ instead of Ax, if A is an $n \times n$ matrix.]

(i) Let $b \in \mathbb{R}^n$ and consider

$$R \circ T_b(x) = R(x + b) = R(x) + R(b)$$

$$= T_a \circ R(x), \tag{P4.5}$$

where $a = R(b)$. Hence, $b = R^{-1}(a)$. In conclusion,

$$T_a \circ R = R \circ T_b. \tag{P4.5a}$$

(ii) We show that $M_1 \circ M_2$ is a motion. Then $M_1 = T_a \circ U$ and $M_2 = T_b \circ V$, where U and V are othogonal matrices. Thus,

$$M_1 \circ M_2 = (T_a \circ U) \circ (T_b \circ V)$$

$$= U \circ T_{U^{-1}(a)} \circ T_b \circ V \tag{P4.5b}$$

$$= U \circ T_{U^{-1}(a)+b} \circ V \tag{P4.5c}$$

$$= T_{U^{-1}(U^{-1}(a)+b)} \circ U \circ V, \tag{P4.5d}$$

which is a motion, because $U \circ V = UV$ is an orthogonal matrix.

(iii) We can represent the unity motion as a composition of the unity matrix I (which is orthogonal) and translation T_0.

(iv) We show that for any motion, say M_1, there is a motion M_2 such that $M_1 \circ M_2 = I$. Indeed, let in (P4.5c) $U^{-1}(a) + b = 0$ [implying that $b = -U^{-1}(a)$] and $V = U^{\mathrm{T}}$. Then, from (P4.5c),

$$M_1 \circ M_2 = U \circ T_0 \circ V = I,$$

if $M_2 = T_{-U^{-1}(a)} \circ U^{\mathrm{T}}$. □

4.6 Let f be a homothetic function ($f(x) = \alpha x$) defined in Proposition 4.3, part 1, λ be the Borel - Lebesgue measure on $\mathcal{B}(\mathbb{R}^n)$, λ^* be the Lebesgue outer measure, $\mathcal{L}^*$ be the σ-algebra of Lebesgue measurable sets, and λ_0^* be the Lebesgue measure on $\mathcal{L}^*$. Let μ^* be the outer measure generated by the image measure λf^*, $\mathcal{B}_\mu^*$ be the σ-algebra of all μ^*-measurable sets, and $\mu_0^* = \mathrm{Res}_{\mathcal{B}_\mu^*}\mu^*$. Show that:

a) $\mu^* = |\alpha|^{-n}\lambda^*$ on $\mathcal{P}(\mathbb{R}^n)$.

b) $\mu_0^* = |\alpha|^{-n}\lambda_0^*$ on $\mathcal{B}^*$.

c) $\mathcal{B}_\mu^* = \mathcal{L}^*$.

Solution. a - b) Let μ_1^* be the outer measure generated by Borel measure λf^* and μ_2^* be the outer measure generated by Borel measure $\frac{1}{|\alpha|^n}\lambda$. Then, the outer measures μ_1^*, μ_2^*, and $\frac{1}{|\alpha|^n}\lambda^*$ coincide on $\mathcal{B}$. Indeed, $\mu_i = \mu_i^*$ on $\mathcal{B}$ and $\mu_1 = \mu_2$. In addition,

$$\mu_2 = \tfrac{1}{|\alpha|^n}\lambda = \tfrac{1}{|\alpha|^n}\lambda^* = \mu_2^* \text{ on } \mathcal{B}, \text{ as } \lambda = \lambda^* \text{ on } \mathcal{B}.$$

According to Remark 3.2(i) [Theorem 3.1 states that the Lebesgue content on $\mathcal{R}(\mathcal{S})$ in $\mathbb{R}^n$ is σ-additive. This, obviously, implies that the Lebesgue elementary content is also σ-additive on $\mathcal{S}$.],

$$\mu_1^* = \mu_2^* = \tfrac{1}{|\alpha|^n}\lambda^* \text{ on } \mathcal{P}(\mathbb{R}^n)$$

and their restrictions

$$\mu_{10}^* = \mu_{20}^* = \tfrac{1}{|\alpha|^n}\lambda_0^* \text{ on } \mathcal{B}^*.$$

c) Denote $\mu^* = \frac{1}{|\alpha|^n}\lambda^*$, which by (a-b), is also the outer measure produced by $(\mathcal{B}, \lambda f^*)$. Then, for each $Q \in \mathcal{P}(\mathbb{R}^n)$, the equation

$$\mu^*(Q) = \mu^*(Q \cap B) + \mu^*(Q \cap B^c)$$

holds true if and only if

$$\lambda^*(Q) = \lambda^*(Q \cap B) + \lambda^*(Q \cap B^c)$$

does. Therefore, $B \in \mathcal{B}^*$ if and only if $B \in \mathcal{L}^*$. □

4.8 Show that the Lebesgue measure λ_0^* on $\mathcal{L}^*$ is translation-invariant.

Solution.

(i) First we need to show that for each $\mathcal{L}^*$-measurable set A,

$$A + x \in \mathcal{L}^*.$$

From Remarks and Definitions 3.2(ii) [or more specifically, from Theorem 2.19(ii)], $(\mathbb{R}^n, \mathcal{L}^*, \lambda_0^*)$ is complete and coincides with the completion of $(\mathbb{R}^n, \mathcal{B}, \lambda)$. Hence, each $A \in \mathcal{L}^*$ can be represented as $A = B \cup M,$ where $B \in \mathcal{B}$ and M is negligible and, consequently,

$$A + x = (B + x) \cup (M + x)$$

To prove that $A + x \in \mathcal{L}^*$ it is sufficient to show that $M + x$ is negligible, as by Problem 3.10, $B + x \in \mathcal{B}$. Because M is negligible, there is a Borel measurable superset, say N, such that $\lambda(N) = 0$. Clearly, $N + x$ is again Borel and because Lebesgue-Borel measure λ is translation-invariant (Proposition 4.3), $\lambda(N + x) = 0$. Obviously, $N + x$ is a superset of $M + x$. So, we are done with (i).

(ii) If $\overline{\lambda}$ is the completion of $\lambda,$ then $\overline{\lambda}(A + x) = \lambda(B + x)$ and by the translation-invariance of $\lambda,$ the latter is equal to $\lambda(B)$. Because $\overline{\lambda} = \lambda_0^*,$ we have that

$$\lambda_0^*(A + x) = \lambda(B) = \lambda_0^*(B \cup M) = \lambda_0^*(A).$$ □

4.10 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and X_i be a (Ω, Σ)-(Ω_i, Σ_i) r.v., $i = 1, \ldots, n$. Show that $X_1, \ldots, X_n$ are independent if and only if σ-algebras Σ_i's have $\cap$-stable generators $\mathcal{G}_i$

such that the system $\{X_1^{**}(\mathcal{G}_1), \ldots, X_n^{**}(\mathcal{G}_n)\}$ of families of events is independent.

Solution. (i) Suppose $\{X_1^{**}(\mathcal{G}_1), \ldots, X_n^{**}(\mathcal{G}_n)\}$ are independent. Then, if $\mathcal{G}_i$ is a generator of Σ_i, $X_i^{**}(\mathcal{G}_i)$ is a generator of $X_i^{**}(\Sigma_i)$. Indeed, by Problem 3.3, Chapter 4, if $f : \Omega \to \Omega'$ is a function and $\mathcal{G}' \subseteq \mathcal{P}(\Omega')$, then

$$f^{**}(\Sigma(\mathcal{G}')) = \Sigma(f^{**}(\mathcal{G}')),$$

or in the form

$$\Sigma(X_i^{**}(\mathcal{G}_i)) = X_i^{**}(\Sigma_i) = X_i^{**}(\Sigma(\mathcal{G}_i)).$$

Furthermore, it is easily seen that with $\mathcal{G}_i$'s being $\cap$-stable, so are $X_i^{**}(\mathcal{G}_i)$. Thus, by Problem 1.22, the system of σ-algebras $\{X_1^{**}(\Sigma_1), \ldots, X_n^{**}(\Sigma_n)\}$ is independent and so are the r.v.'s $X_1, \ldots, X_n$. We only notice that Problem 1.22 required that the generators $\mathcal{G}_i$'s be semi-rings, but in the proof of Problem 1.22 $\cap$-stability was all we needed.

(ii) The converse is trivial. If $X_1, \ldots, X_n$ are independent, then the σ-algebras $\{X_1^{**}(\Sigma_1), \ldots, X_n^{**}(\Sigma_n)\}$ can serve as $\cap$-stable generators of themselves. □

5.4 Show that if $f^3 \in \mathcal{C}^{-1}$, then $f \in \mathcal{C}^{-1}$. Show that if $f^2 \in \mathcal{C}^{-1}$ then f need not be in $\mathcal{C}^{-1}$.

Solution. (i) f^3 is measurable if and only if $\{f^3 \leq a\}$ for all real values of a. But then, $\{f^3 \leq a\} = \left\{f \leq a^{1/3}\right\}$ if we consider only real values of a.

(ii) Let A be a nonmeasurable set and $f = \mathbf{1}_A - \mathbf{1}_{A^c}$. Then, f is a nonmeasurable function considering that $f : \Omega \to \{1, -1\}$. However, $f^2 = \mathbf{1}_A + \mathbf{1}_{A^c} - 0 = \mathbf{1}$ is a measurable function with respect to the trivial σ-algebra $\{\Omega, \emptyset\}$. □

5.8 Show that if f has a derivative at each point of $\mathbb{R}$, then this derivative is Borel-measurable.

Solution. If f has a derivative on $\mathbb{R}$, then for each x,

$$f'(x) = \lim_{n \to \infty} \frac{f(x+h_n)-f(x)}{h_n},$$

where $\{h_n\}$ is a sequence convergent to zero. Denoting

$$f_n(x) = \frac{f(x+h_n)-f(x)}{h_n}$$

we see that for each n, f_n is Borel-measurable. This is because f is differentiable and therefore continuous. Now the statement is due to Proposition 5.6(vi). □

5.9 Show that the space $(\mathcal{C}^{-1}, \tau_P)$ is Hausdorff.

Solution. First note that $\mathcal{C}^{-1}$ is a subset of $\overline{\mathbb{R}}^{\Omega}$ and because $(\overline{\mathbb{R}}, \overline{\tau}_e)$ is Hausdorff, the Tychonov product space $\left(\overline{\mathbb{R}}^{\Omega}, \tau_P\right)$ is Hausdorff, as we know from Problem 5.17, Chapter 3. The statement follows from Problem 6.12, Chapter 3, according to which Hausdorffness is hereditary. □

5.10 Let (Ω, Σ, μ) be a complete measure space and let $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Suppose that $g\colon \Omega \longrightarrow \overline{\mathbb{R}}$ is an extended, real-valued function. Show that if $g = f \pmod{\mu}$, then $g \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$.

Solution. The assumption $g = f \pmod{\mu}$ implies that $N = \{f \neq g\}$ is a negligible set. Because the measure space (Ω, Σ, μ) is complete, both N and $N^c \in \Sigma$. Obviously, for any Borel set B,

$$g^*(B) \cap N^c = f^*(B) \cap N^c. \tag{P5.10}$$

Because N is negligible, so is $g^*(B) \cap N$ as a subset of N. Because (Ω, Σ, μ) is complete, $g^*(B) \cap N$ is measurable. From (P5.10) it follows that $g^*(B) \cap N^c$ is also measurable. Hence,

$$g^*(B) = [g^*(B) \cap N] \cup [g(B^*) \cap N^c]$$

is measurable. Therefore, $g^{**}(\mathcal{B}(\overline{\mathbb{R}})) \subseteq \Sigma$ implying that $g \in \mathcal{C}^{-1}(\Omega, \Sigma)$. □

5.11 Let (Ω, Σ, μ) be a complete measure space and let $\{f_n\} \subseteq \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Suppose that $\lim_{n\to\infty} f_n$ exists and $f_n \to f$ pointwise μ-a.e. on Ω, where f is an extended, real-valued function. Show that $f \in \mathcal{C}^{-1}$.

Solution. By Proposition 5.6(vi),

$$\lim_{n\to\infty} f_n \in \mathcal{C}^{-1}(\Omega,\Sigma).$$

Because $\lim_{n\to\infty} f_n = f \pmod{\mu}$, due to Problem 5.10, we have $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$. □

5.12 Let Q be a non-Borel subset of $\mathbb{R}$ (such as one in Problem 3.16) and let C denote the Cantor ternary set. Define the function

$$f(x) = \mathbf{1}_{Q\cap C}(x)\sin x + \mathbf{1}_{(Q\cap C)^c}(x)x^2.$$

Is f Lebesgue measurable, that is, $f \in \mathcal{C}^{-1}(\mathbb{R}, \mathcal{L}^*, \lambda_0^*)$?

Solution. Because $Q \cap C$ is negligible, $f(x) = x^2 \pmod{\lambda}$. Because the Lebesgue measure space $(\mathbb{R}, \mathcal{L}^*, \lambda_0^*)$ is complete, by Problem 5.10, f is Lebesgue-measurable. □

5.13 We knew from probability theory that r.v.'s $X_1, \ldots, X_n$ on a probability space $(\Omega, \Sigma, \mathbb{P})$ and valued in $\overline{\mathbb{R}}$ (each), endowed with the Borel σ-algebra $\mathcal{B}(\overline{\mathbb{R}})$, are independent if and only if their joint probability distribution function (commonly abbre-

viated as the PDF) $\mathbb{P}\{X_1 \le x_1, \dots, X_n \le x_n\}$ equals the product $\prod_{i=1}^{n} \mathbb{P}\{X_i \le x_i\}$ of their marginal PDF's. Prove that this is true using the original definition in Example 4.5(iii).

Solution. (i) Suppose

$$\mathbb{P}\{X_1 \le x_1, \dots, X_n \le x_n\} = \prod_{i=1}^{n} \mathbb{P}\{X_i \le x_i\}$$

holds true. From Proposition 5.2, we learned that the collection of sets $\mathcal{G} = \{[-\infty, a] : a \in \mathbb{R}\}$ is a generator of the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. Clearly, $\mathcal{G}$ is $\cap$-stable. Therefore, by Problem 4.10, $X_1, \dots, X_n$ are independent if and only if $\Sigma_i = \mathcal{B}(\mathbb{R})$ has an $\cap$-stable generator such that the system $\{X_1^{**}(\mathcal{G}_1), \dots, X_n^{**}(\mathcal{G}_n)\}$ of families of events (with $\mathcal{G}_i = \mathcal{G}$) is independent. The latter means (see also Problem 4.11) that

$$\mathbb{P}\Big[\{X_1 \in [-\infty, x_1]\} \cap \cdots \cap \{X_n \in [-\infty, x_n]\}\Big]$$

$$= \prod_{i=1}^{n} \mathbb{P}\{X_i \in [-\infty, x_i]\}$$

is a sufficient condition that the system of σ-algebras $\{X_1^{**}(\Sigma_1), \dots, X_n^{**}(\Sigma_n)\}$ is independent and so are the r.v.'s $X_1, \dots, X_n$.

(ii) The converse is due to Problem 4.10. □

6.2 Let Ω be an uncountable set and let $\Sigma = \{A \subseteq \Omega :\ A \text{ or } A^c \text{ is at most countable}\}$. Show that $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$ if and only if f is constant everywhere except on an at most countable subset of Ω.

Solution. (i) Let f be measurable.

Case (a) $f = \mathbf{1}_A$. Then, $A \in \Sigma$ and thus either A or $A^c \preceq \mathbb{N}$. Either way, f meets the assertion.

Case (b) $f = \sum_{i=1}^{m} c_i \mathbf{1}_{A_i}$. Then, at least one of the A_i's is uncountable on which f is a constant.

Case (c) $f \in \mathcal{C}_+^{-1}$. Then, there is a monotone sequence $\{g_n\}$ of simple functions such that $f = \sup g_n$. Suppose B_n is an uncountable set on which $g_n = b_n$. Obviously, $\{b_n\}$ is a monotone non-decreasing sequence whose limit equals f on some set. We claim that surely $f(\omega) = \sup b_n$ on $\bigcap_{n=1}^{\infty} B_n$. Indeed, suppose we have

$$g_1 \leq g_2 \text{ and } g_1 = b_1 \text{ on } B_1 \text{ and } g_2 = b_2 \text{ on } B_2.$$

Clearly,

$$g_2(\omega) = g_1(\omega) \vee g_2(\omega) = b_2 \text{ on } B_1 \cap B_2.$$

Now, obviously,

$$\left(\bigcap_{n=1}^{\infty} B_n\right)^c = \bigcup_{n=1}^{\infty} B_n^c \preceq \mathbb{N}.$$

Thus, $B := \bigcap_{n=1}^{\infty} B_n$ is uncountable. Hence, $f = b = \sup b_n$ on B.

Case (d) $f = f^+ - f^-$. Suppose $f^+ = b$ on B and $f^- = c$ on C. Then, because $B^c \cup C^c \preceq \mathbb{N}$, we have that $B \cap C$ is uncountable and thus, $f = b - c$ on $B \cap C$.

(ii) The converse. Let $f = c$ on some A such that A is uncountable and $A^c \preceq \mathbb{N}$. Consider the set

$$\{f \geq c\} = \{f = c\} + \{f > c\} = A + B.$$

Obviously,

$$\{f < c\} = A^c \cap B^c \preceq \mathbb{N}.$$

Thus, $\{f \geq c\} \in \Sigma$ and so is $\{f < c\}$. Let $\alpha \leq c$. Then,

$$\{f < \alpha\} \subseteq \{f < c\} \preceq \mathbb{N},$$

whereas

$$\{f \geq c\} \subseteq \{f \geq \alpha\}$$

is uncountable. Thus, for all $\alpha \leq c$, $\{f < \alpha\}$ and $\{f \geq \alpha\}$ are measurable. It is easy to see that these sets are measurable for $a > c$ as well. □

6.5 Theorem 6.4 states that for each $f \in \mathcal{C}_+^{-1}$, there is a monotone nondecreasing sequence $\{s_n\}$ of nonnegative simple functions (from Ψ_+) such that $\sup\{s_n\} = f$. In Remark 6.5, it is shown that unless $\mathcal{C}_+^{-1}$ is first countable in the sense of τ_P (which in most applications it is not), $\mathcal{C}_+^{-1}$ is a proper subset of $\overline{\Psi}_+$ (the τ_P-closure of Ψ_+). Does it mean that the converse of Theorem 6.4 does not hold [i.e., that, although $\sup\{s_n\} \in \overline{\Psi}_+$ of some sequence $\{s_n\}$ of simple functions, it need not belong to $\mathcal{C}_+^{-1}$]? Analyze and explain this.

Solution. By Theorem 5.6(vi), $\mathcal{C}^{-1}$ is closed with respect to all pointwise limits. In other words, $(\mathcal{C}^{-1}, \tau_P)$ contains all of its sequential limit points, as does the relative topology $(\mathcal{C}_+^{-1}, \tau_P \cap \mathcal{C}_+^{-1})$. Therefore, all monotone nondecreasing sequences of simple functions (which are clearly pointwise convergent) will have their limits in $\mathcal{C}_+^{-1}$.

However, $(\mathcal{C}_+^{-1}, \tau_P \cap \mathcal{C}_+^{-1})$ does not contain its closure points, because as argued in Remark 6.4, for most cases, it remains a proper subset of $\overline{\Psi}_+$. A topological space, as we learned from Remark 1.15 and Examples 1.16, can contain all of its sequential limit points and yet not being closed. In conclusion, Theorem 6.4 can be extended by the inclusion of the phrase: "Conversely, every limit of a monotone nondecreasing sequence of simple functions belongs to $\mathcal{C}_+^{-1}$." □

Chapter 6. Integration in Abstract Spaces

1.3 Let (Ω, Σ, μ) be a measure space and $c > 0$. Show that for each $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$,

$$\int f d(c\mu) = c \int f d\mu.$$

Solution. First observe that $c\mu$ is a measure on (Ω, Σ). Let

$$f = \sum_{k=1}^{n} a_k \mathbf{1}_{A_k} \in \Psi_+(\Omega, \Sigma).$$

Then,

$$\int f d(c\mu) = \sum_{k=1}^{n} a_k c\mu(A_k) = c \int f d\mu.$$

Let $f \in \mathcal{C}_+^{-1}$. Then, there is $\{s_n\} \uparrow \subseteq \Psi_+(\Omega, \Sigma)$ such that $f = \sup s_n$. Hence,

$$\int f d(c\mu) = \sup \int s_n d(c\mu) = \sup\{c \int s_n d\mu\} = c \int f d\mu.$$

Finally, for $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$, the proof is routine. □

1.5 Let $f \in \mathcal{C}^{-1}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ and suppose that there are two non-negative integrable functions u and v such that $f = u - v$. Show that $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ and that

$$\int f d\mu = \int u d\mu - \int v d\mu.$$

Solution. (i) Let $f = f^+ - f^-$. Then,

$$f^+ - f^- = u - v \Rightarrow f^+ + v = u + f^-. \tag{P1.5}$$

Furthermore,

$$f = u - v \leq u \leq u + v =: h$$

and

$$-f = v - u \leq v \leq u + v = h$$

Then, from

$$-h \leq f \leq h \;\Rightarrow |f| \leq h$$

and because by Proposition 1.3(ii) h is integrable, so is $|f|$, by Proposition 1.3(iii). The latter means that $f \in L^1$.

(ii) Now, because by Proposition 1.8, L^1_+ is a semivector space, we have:

$$\int (f^+ + v)d\mu = \int f^+ d\mu + \int v d\mu$$

$$\int (u + f^-)d\mu = \int u d\mu + \int f^- d\mu.$$

Because of (P1.5), the left-hand sides of the last two equations are identical; so are the right-hand sides and thus we get

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu = \int u d\mu - \int v d\mu. \qquad \square$$

1.7 Show that $(L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}), \mathbb{R}, +, \cdot, \preceq)$ is a Riesz space.

Solution. Because in light of Problem 1.6, $(L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}), \mathbb{R}, +, \cdot, \preceq)$ is a partially ordered vector space [see Definition 7.7(i)], it remains to show that $f, g \in L^1$ implies that $f \vee g \in L^1$ and $f \wedge g \in L^1$. Clearly,

$$|f \vee g| \preceq |f| + |g| \quad \text{and} \quad |f \wedge g| \preceq |f| + |g|$$

that yields the assertion. $\square$

1.9 Let (Ω, Σ, μ) be a measure space and let $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$. Show that $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ if and only if there exists $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$ such that $|f| \leq g$.

Solution.

(i) If such a $g \in L^1(\Omega, \Sigma, \mu)$ exists with $|f| < g$, then, by monotonicity of the integral (Proposition 1.12), $|f| \in L^1(\Omega, \Sigma, \mu)$ and hence $f \in L^1(\Omega, \Sigma, \mu)$, as $f \leq |f|$.

(ii) $f \in L^1(\Omega, \Sigma, \mu)$ if and only if $|f| \in L^1(\Omega, \Sigma, \mu)$, as noticed in Definition 1.10(i). Hence $|f|$ can be taken as g. □

1.11 Show that if $f \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ then $f \in [\mathbf{0}]_\mu$ yields that $\int f d\mu = 0$. Does the converse hold true?

Solution. By Problem 6.4, Chapter 5, $f \in [\mathbf{0}]_\mu \Rightarrow f^+, f^- \in [\mathbf{0}]^+_\mu$. By Lemma 1.15, both $\int f^+ d\mu$ and $\int f^- d\mu$ equal zero. Thus, $\int f d\mu = 0$. However, if $\int f^+ d\mu$ and $\int f^- d\mu$ are equal (such as if $f(x) = \sin x$ on $[0, 2\pi]$), then $\int f d\mu = 0$, whereas f is not. □

1.13 Generalize Proposition 1.16. *Let* (Ω, Σ, μ) *be a measure space and let* $f, g \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ *such that* $f = g$ *(mod* μ*) and* $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. *Then*

$$\int f d\mu = \int g d\mu$$

and $g \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$.

Solution. Suppose that $f = g \pmod{\mu}$ and that $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$. By Problem 6.3, Chapter 5,

$$f^+ = g^+ \pmod{\mu} \text{ and } f^- = g^- \pmod{\mu}.$$

By Proposition 1.16 [*If* (Ω, Σ, μ) *is a measure space and* $f, g \in \mathcal{C}^{-1}_+(\Omega, \Sigma; \overline{\mathbb{R}})$ *such that* $f = g \pmod{\mu}$, *then* $\int f d\mu = \int g d\mu$],

$$\int f^+ d\mu = \int g^+ d\mu \text{ and } \int f^- d\mu = \int g^- d\mu. \qquad \square$$

1.15 Show that for $f \in \mathcal{C}^{-1}(\Omega, \Sigma)$, $\int\limits_A f d\mu = 0$ for each $A \in \Sigma$ if and only if $f \in [\mathbf{0}]_\mu$. Does this assertion generalize Lemma 1.15?

Solution.

(i) If $f \in [\mathbf{0}]_\mu$, then for each $A \in \Sigma$, $1_A f \in [0]_\mu$, and therefore,

$$\int_A f d\mu = 0$$

follows from Problem 1.12.

(ii) Let $\int_A f d\mu = 0$ for each $A \in \Sigma$. Denote

$$M = \{f > 0\} \text{ and } N = \{f < 0\}.$$

Clearly, $M, N \in \Sigma$. By the above assumption,

$$\int f 1_M d\mu = \int (-f) 1_N d\mu = 0.$$

Because both integrands in the last equation are elements of $\mathcal{C}_+^{-1}(\Omega, \Sigma)$, by Lemma 1.15,

$$f\mathbf{1}_M, (-f)\mathbf{1}_N \in [\mathbf{0}]_\mu^+.$$

Hence,

$$\mu(M) = \mu(N) = \mu(M + N) = \mu(\{f \neq 0\}) = 0.$$

(iii) No, this assertion does not generalize Lemma 1.15, because it requires that $\int_A f d\mu = 0$ for each $A \in \Sigma$, whereas Lemma 1.15 does not. □

1.17 Let π_α be a Poisson measure and let $f \in \mathcal{C}_+^{-1}(\mathbb{R}, \mathcal{B}; \overline{\mathbb{R}})$. Show that

$$\int f d\pi_\alpha = \sum_{n=0}^{\infty} e^{-\alpha} \frac{\alpha^n}{n!} f(n).$$

Solution. Using Example 1.9 (iv) and Problem 1.3 we have the validity of the assertion by setting $\mu_n = e^{-\alpha} \frac{\alpha^n}{n!} \varepsilon_n, n \in \mathbb{N}_0$. □

1.20 Let μ and ν be two measures on (Ω, Σ) such that $\mu \leq \nu$. Show that for each $f \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}})$, the integral $\int f d(\nu - \mu)$ makes sense, $f \in L^1(\Omega, \Sigma, \nu - \mu; \overline{\mathbb{R}})$, and that

$$\int f d(\nu - \mu) = \int f d\nu - \int f d\mu.$$

Solution. Obviously, $\nu - \mu$ is a measure on Σ and for each $A \in \Sigma$, it holds true that

$$(\nu - \mu)(A) = \nu(A) - \mu(A).$$

Let $s = \sum_{k=1}^n \alpha_k \mathbf{1}_{A_k} \in \Psi_+(\Omega, \Sigma)$. Then

$$\begin{aligned} \int s d(\nu - \mu) &= \textstyle\sum_{k=1}^n \alpha_k (\nu - \mu)(A_k) \\ &= \textstyle\sum_{k=1}^n \alpha_k \nu(A_k) - \sum_{k=1}^n \alpha_k \mu(A_k), \end{aligned} \tag{P1.20}$$

which yields the statement for integrable simple functions. Note that $s \in L^1(\Omega, \Sigma, \nu - \mu)$, as follows from (P1.20), because $s \in L^1(\Omega, \Sigma, \mu) \cap L^1(\Omega, \Sigma, \nu)$. If f is nonnegative and $\{s_n\} \uparrow f$, then from (P1.20),

$$\begin{aligned} \int f d(\nu - \mu) &= \sup \int s_n d(\nu - \mu) \\ &= \sup \Big(\textstyle\sum_{k=1}^{k_n} \alpha_k^{(n)} \nu(A_{kn}) - \sum_{k=1}^{k_n} \alpha_k^{(n)} \mu(A_{kn}) \Big) \\ &= \int f d\nu - \int f d\mu. \end{aligned} \tag{P1.20a}$$

Again, (P1.20a) yields that $f \in L^1(\Omega, \Sigma, \nu - \mu)$. Finally, if $f \in L^1(\Omega, \Sigma, \mu) \cap L^1(\Omega, \Sigma, \nu)$, then $f = f^+ - f^-$ and

$$\begin{aligned} \int f d(\nu - \mu) &= \int f^+ d(\nu - \mu) - \int f^- d(\nu - \mu) \\ &= \int f^+ d\nu - \int f^- d\nu - (\int f^+ d\mu - \int f^- d\mu) \end{aligned}$$

and the statement follows. □

1.22 Prove Proposition 1.18. *Show that if* $f, g \in \mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})$ *and* f *or* $g \in L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$, *then*

$$\int_A f d\mu = \int_A g d\mu \ \textit{for each } A \in \Sigma \tag{P1.22}$$

yields that $f = g$ (*mod* μ).

Solution. Let $N = \{f > g\}$. Then, the function $(f - g)\mathbf{1}_N \in \mathcal{C}_+^{-1}$. From (P1.22) it follows that $g \in L^1(\Omega, \Sigma, \mu)$ if f does. Equation (P1.22) and the linearity of $L^1(\Omega, \Sigma, \mu)$ (see Problem 1.8) imply that

$$\int (f - g)\mathbf{1}_N d\mu = 0.$$

By Lemma 1.15, $(f - g)\mathbf{1}_N \in [\mathbf{0}]_\mu^+$. Similarly, $(f - g)\mathbf{1}_M \in [\mathbf{0}]_\mu^+$, where $M = \{g > f\}$. Therefore,

$$(f - g)\mathbf{1}_{\{f \neq g\}} \in [\mathbf{0}]_\mu^+$$

from which it follows that $\{f \neq g\}$ is a μ-null set or that $f - g \in [\mathbf{0}]_\mu$. □

1.25 Let $g \in \mathcal{C}^{-1}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ such that $\int_A g\, d\mu \geq 0$ for all $A \in \Sigma$. Show that $g \geq 0$ μ-a.e.

Solution. Let $A = B \cap \{g < 0\}$, B an arbitrary measurable set, and let $h = g\,\mathbf{1}_{\{g<0\}}$. Then we have $\int h\mathbf{1}_B d\mu \geq 0$, whereas $h \leq 0$ on Ω. The latter yields that $\int h\mathbf{1}_B d\mu \leq 0$. Comparing the two inequalities we conclude that

$$\int h\mathbf{1}_B d\mu = 0 \ \text{ for every measurable set } B.$$

By Problem 1.15, then $h = 0$ (mod μ) implying that $g\,\mathbf{1}_{\{g<0\}} = 0$ μ-a.e. It obviously implies that $\{g < 0\}$ is a μ-null set. Consequently, $g \geq 0$ μ-a.e. □

2.2 Generalize the monotone convergence theorem: Let $\{f_n\}\uparrow \subseteq \mathcal{C}^{-1}(\Omega, \Sigma)$ and $g \in \mathcal{C}^{-1}(\Omega, \Sigma)$ such that $f_n \geq g$ $f$or all n and suppose that $\int g\, d\mu > -\infty$. Prove that

$$\sup\{\int f_n\, d\mu\} = \int \sup\{f_n\}\, d\mu.$$

Solution.

(i) Suppose $\int g < \infty$, which, along with $\int g\, d\mu > -\infty$ means that $g \in L^1(\Omega, \Sigma, \mu)$. Therefore,

$$f_n \geq g \;\Rightarrow\; f_n - g \geq 0,\; \{f_n - g\} \subseteq \mathcal{C}_+^{-1},$$

and $\{f_n - g\}$ is monotone nondecreasing. Hence, we can apply to $\{f_n - g\}$ the monotone convergence theorem 2.1:

$$\sup\int (f_n - g) = \int \sup(f_n - g).$$

The last equation yields

$$\sup\int f_n - \int g = \int \sup f_n - \int g,$$

which in turn reduces to

$$\sup\int f_n = \int \sup f_n,$$

because by the above assumption $g \in L^1(\Omega, \Sigma, \mu)$.

(ii) Let $\int g = +\infty$. Then, from $f_n \geq g$ it follows that

$$\int f_n \geq \int g = +\infty \;\Rightarrow\; \sup\int f_n = +\infty.$$

On the other hand, $\sup f_n \geq f_n \geq g \;\Rightarrow\; \int \sup f_n \geq \int g = +\infty$. □

2.4 Let $\{f_n\}\downarrow \subseteq \mathcal{C}^{-1}$ and $g \in \mathcal{C}^{-1}$ such that $f_n \leq g$ for all n. If $\int g\, d\mu < \infty$, show that

$$\inf\{\int f_n\, d\mu\} = \int \inf\{f_n\}\, d\mu.$$

Solution. Apply the generalized monotone convergence theorem to $\{-f_n\}\uparrow$ and $-g$. □

2.6 Show that if $f_n \to f$ in mean then

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Solution.

$$\begin{aligned} \left| \int (f_n - f) d\mu \right| &= \left| \int f_n d\mu - \int f d\mu \right| \\ &\leq \int |f_n - f| \, d\mu = : a_n. \end{aligned}$$

Because $a_n \to 0$, for each ε, there is an N such that for all $n \geq N$, $a_n < \varepsilon$. □

2.9 Let

$$f_n(x) = \begin{cases} n, & 0 < x < \frac{1}{n} \\ 0, & \frac{1}{n} \leq x < \infty. \end{cases}$$

Explain why

$$\int \lim_{n\to\infty} f_n \, \lambda(dx) \neq \lim_{n\to\infty} \int f_n \, \lambda(dx).$$

Solution. (i) Clearly,

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} \infty, & x = 0 \\ 0, & x > 0 \end{cases}.$$

Therefore, f is Borel-measurable and $f_n \to \mathbf{0}$ λ-a.e. pointwise on $\mathbb{R}_+$. Because $f = \mathbf{0}$ λ-a.e.,

$$\int_{\mathbb{R}_+} f d\lambda = \int_{\mathbb{R}_+} \mathbf{0} \, d\lambda = 0.$$

On the other hand, $\int_{\mathbb{R}_+} f_n d\lambda = 1$ for all $n = 1, 2, \ldots$.

(ii) Now the reason why the integrals are not equal is because we cannot use the Lebesgue dominated convergence theorem. The latter does not apply because there is obviously no dominant integrable function $g \geq 0$ for the sequence $\{f_n\}$ required in the Lebesgue dominated convergence theorem.

(iii) Here is yet more evidence why the Lebesgue dominated convergence theorem does not apply, because if it would, then by the Lebesgue dominated convergence theorem part (iii),

$$\lim_{n\to\infty} \int |f - f_n| d\lambda = \lim_{n\to\infty} \int |\mathbf{0} - f_n| d\lambda$$

should equal zero, which is not true because it equals 1. Thus, although $\{f_n\}$ converges to $\mathbf{0}$ pointwise almost everywhere on $\mathbb{R}_+$, it does not converge to $\mathbf{0}$ in mean, that is in the L^1-norm. Clearly, there is no other function to which $\{f_n\}$ converges both pointwise a.e. and in mean. □

2.10 Let

$$f_n(x) = \begin{cases} n^2 x, & 0 \leq x \leq \frac{1}{n} \\ -n^2\big(x - \frac{2}{n}\big), & \frac{1}{n} \leq x \leq \frac{2}{n} \\ 0, & x \geq \frac{2}{n}. \end{cases}$$

Show that

$$\int \underline{\lim}\, f_n \, d\lambda < \underline{\lim} \int f_n \, d\lambda.$$

Explain why the Lebesgue dominated convergence theorem cannot be applied.

Solution. It is easily seen that the graph of f_n represents a triangle on interval $\big[0, \frac{2}{n}\big]$, symmetric about $\frac{1}{n}$, with a vertex at $\big(\frac{1}{n}, n\big)$, and that f_n is zero on $\big[0, \frac{2}{n}\big]^c$. Therefore, the function is continuous and integrable. Clearly, $f_n \to \mathbf{0}$ pointwise λ-a.e. More specifically, f_n converges pointwise to function

$$f(x) = \begin{cases} \infty, & x = 0 \\ 0, & x > 0 \end{cases}.$$

Thus f is zero λ-a.e. on $\mathbb{R}_+$. Obviously, the Lebesgue integral of f is zero. On the other hand, evaluation of

$$\int_0^\infty f_n(x)\lambda(dx) = (\mathrm{R})\int_0^\infty f_n(x)dx$$

gives $\frac{1}{2}+\frac{1}{2n^2} \to \frac{1}{2}$ pointwise everywhere on $\mathbb{R}$. This validates the inequality.

We remark that treating f as the Dirac delta requires an additional assumption imposed on f that

$$\int_{-\infty}^\infty f(x)g(x)dx = g(0),$$

which in particular, would imply that $\int_0^\infty f(x)dx = \frac{1}{2}$ and thus the equality of the above two integrals. However, the latter would contradict the Lebesgue integration and yield a different result. The Dirac delta is a so-called *generalized function* and its treatment falls into a different area of analysis known as "generalized functions." The reader should be aware of the seeming discrepancy.

As in Problem 2.9, we cannot use the Lebesgue dominated convergence theorem, because there is no dominant integrable function $g \geq 0$ for the sequence $\{f_n\}$ required so in the Lebesgue dominated convergence theorem.

2.12 Give an example of a monotone nonincreasing sequence of measures convergent to a set function μ setwise such that μ is not a measure.

Solution. Let λ be the Lebesgue measure on $(\mathbb{R}, \mathcal{B})$. Define $\mu_n = \frac{1}{n}\lambda$. We show that $\lim_{n\to\infty} \mu_n = \mu$ is not σ-additive. Take $A_1 = [0, 1)$, $A_2 = [1, 2), \ldots$ as a decomposition of $\mathbb{R}_+ = [0, \infty)$. Because $\mu_n(\mathbb{R}_+) = \infty$ for each n, it follows that $\mu(\mathbb{R}_+) = \infty$. Now as

$\mu_n(A_k) = \frac{1}{n}$, it follows that $\mu(A_k) = 0$ for all k. Hence, μ is not σ-additive. □

2.14 Let $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ and suppose that $\{\Omega_n\}$ is a countable measurable partition of Ω. Show that

$$\int f d\mu = \sum_{n=1}^{\infty} \int_{\Omega_n} f d\mu.$$

Solution. We start with

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu$$

and then have

$$\int f^+ d\mu = \int f^+ \mathbf{1}_{\sum_{n=1}^{\infty} \Omega_n} d\mu$$

by Beppo Levi

$$= \sum_{n=1}^{\infty} \int_{\Omega_n} f^+ d\mu.$$

So, also

$$\int f^- d\mu = \sum_{n=1}^{\infty} \int_{\Omega_n} f^- d\mu.$$

Thus,

$$\int f d\mu = \sum_{n=1}^{\infty} \int_{\Omega_n} f^+ d\mu - \sum_{n=1}^{\infty} \int_{\Omega_n} f^- d\mu.$$

Because at least one of the series is finite, we have

$$\int f d\mu = \sum_{n=1}^{\infty} \int_{\Omega_n} f d\mu.$$ □

3.2 In Example 3.4, we showed that the Dirichlet function f on [0,1] is Lebesgue integrable, but not Riemann integrable. Because the rationals have the Lebesgue measure 0, the function f is equal to 0 (a constant) for λ-almost all points on $[0, 1]$, and therefore, it is continuous almost everywhere on

$[0,1]$. By Theorem 3.5, f must be Riemann integrable. This is just the opposite of the result of Example 3.4. What is wrong with this reasoning?

Solution. It looks as if f is continuous almost everywhere, in fact, the restriction of f on $[0,1]\backslash\mathbb{Q}$ is continuous. However, f is clearly discontinuous at every point if defined on $[0,1]$. In other words, the set of discontinuities of f is not a λ-null set. □

3.3 Is the function $f(x)=\frac{1}{x}$ on $[0,1]$ Borel-measurable and λ-integrable? Is $f\in\mathbb{L}([0,1],\mathcal{B}([0,1],\lambda))$?

Solution. (i) Let

$$f_n(x)=\begin{cases} n, & x\in\left[0,\frac{1}{n}\right] \\ \frac{1}{x}, & x\in\left[\frac{1}{n},1\right] \end{cases}.$$

Obviously, f_n is continuous and bounded on $[0,1]$ and therefore, Borel. Because the sequence $\{f_n\}$ pointwise converges to f, the latter is also Borel.

(ii) Clearly, the sequence $\{f_n\}$ is monotone increasing, so that $f=\lim_{n\to\infty}f_n$ in the topology of pointwise convergence. Therefore, we can apply the monotone convergence theorem to get

$$\begin{aligned}(\mathrm{L})\textstyle\int_0^\infty f d\lambda &= \lim_{n\to\infty}(\mathrm{L})\textstyle\int_0^\infty f_n(x)\lambda(dx)\\ &= \lim_{n\to\infty}(\mathrm{R})\textstyle\int_0^\infty f_n(x)dx\\ &= \lim_{n\to\infty}(1+\ln n)=\infty.\end{aligned}$$

(iii) Consequently, $f\notin L^1([0,1],\mathcal{B}([0,1],\lambda))$. However, because f is nonnegative and the integral of f is $+\infty$, it exists and hence $f\in\mathbb{L}([0,1],\mathcal{B}([0,1],\lambda))$. □

3.5 Let $f:[0,1]\to\mathbb{R}$ be defined as

$$f(x) = \begin{cases} (-1)^{n+1}n, & x \in A_n = (\frac{1}{n+1}, \frac{1}{n}],\ n = 1, 2, \dots \\ 0, & x = 0. \end{cases}$$

Show that f is improperly Riemann integrable but not Lebesgue integrable.

Solution.

$$(\mathrm{R})\int_0^1 f(x)dx = \sum_{n=1}^{\infty} \int_{A_n} n(-1)^{n+1} dx$$

$$= \sum_{n=1}^{\infty} (-1)^{n+1} \tfrac{1}{n+1} < \infty,$$

as a Leibnitz series. On the other hand, $(\mathrm{L})\int_0^1 f d\lambda$ is convergent if and only if $(\mathrm{L})\int_0^1 |f| d\lambda$ is. Hence,

$$(\mathrm{L})\int_0^1 |f| d\lambda = \sum_{n=1}^{\infty} (\mathrm{L}) \int_{A_n} n|(-1)^{n+1}|\lambda(dx)$$

$$= \sum_{n=1}^{\infty} (\mathrm{R}) \int_{A_n} n dx = \sum_{n=1}^{\infty} \tfrac{1}{n+1} = \infty. \qquad \square$$

3.7 Investigate if the function $f(x) = \frac{\sin x}{x^a} \mathbf{1}_{\{x \neq 0\}}(x)$ (where $0 < a \leq 1$) is improperly Riemann and Lebesgue integrable over $\mathbb{R}_+$.

Solution. f is improperly integrable by the Dirichlet criterion applied to $f = gh$, where $g(x) = \frac{1}{x^a}$ and $h(x) = \sin x$. By Proposition 3.9, f is Lebesgue integrable if and only if $\mathrm{R} = (\mathrm{R})\int_0^{\infty} \mid f \mid < \infty$. For $a \leq 1$, we have

$$\mathrm{R} = \sum_{n=0}^{\infty} \int_{n\pi}^{(n+1)\pi} \tfrac{|\sin x|}{x^a}\, dx$$

$$= \sum_{n=0}^{\infty} \int_0^{\pi} \tfrac{\sin(t+n\pi)}{(t+n\pi)^a}\, dt$$

$$= \sum_{n=0}^{\infty} \int_0^{\pi} \frac{\sin t}{(t+n\pi)^a}\, dt$$

$$\geq \frac{2}{\pi^a} \sum_{n=1}^{\infty} \frac{1}{(n+1)^a} = \infty,\ 0 < a \leq 1.$$

For $a > 1$ we use the estimate

$$\mathrm{R} = \sum_{n=0}^{\infty} \int_0^{\pi} \frac{\sin t}{(t+n\pi)^a}\, dt \leq \frac{2}{\pi^a} \sum_{n=1}^{\infty} \frac{1}{n^a} < \infty,\ a > 1. \qquad \square$$

3.9 Show that the functional

$$\| f \|_{L^1} = \int_a^b |f| d\lambda$$

defines a seminorm on $L^1([a,b], \mathcal{B}([a,b]), \lambda)$. How can $\| \cdot \|_{L^1}$ be "made" a norm?

Solution. Obviously, all properties of the norm apply except for $\|f\|_{L^1} = 0$ if and only if $f = \mathbf{0}$. However, by Lemma 1.15, $\|f\|_{L^1} = 0$ if and only if $f \in [\mathbf{0}]_\lambda$. Therefore, we can replace $\| \cdot \|_{L^1}$ with the restrictor $\| \cdot \|_{\boldsymbol{L^1}}$ **(bold face)** acting on the quotient space $L^1([a,b], \mathcal{B}([a,b]), \lambda)|_\lambda$ modulo λ and thus the function

$$(L^1([a,b], \mathcal{B}([a,b]), \lambda)|_\lambda, \mathbb{R}_+, \| \cdot \|_{\boldsymbol{L^1}})$$

indeed defines a norm. $\square$

3.11 Show that the space $\mathcal{C}([a,b])$ of all continuous functions on interval $[a,b]$ is dense in $(L^1([a,b], \mathcal{B}([a,b]), \lambda), \| \cdot \|_{L^1})$.

Solution. Let

$$f \in L^1([a,b], \mathcal{B}([a,b]), \lambda).$$

We show that for any $\varepsilon > 0$, there is a continuous function H such that $\| f - H \|_{L^1} < \varepsilon$. For $F \geq 0$ such that $F \in L^1([a,b], \mathcal{B}([a,b]), \lambda)$, there is a monotone increasing sequence

$\{s_n\} \subseteq \mathcal{F}_+$ such that $\sup s_n = F$. Thus, for a given ε, there is an N such that for all $n \geq N$,

$$\|F - s_n\|_{L^1} < \varepsilon.$$

By Problem 3.10, for s_N, there is a continuous function h such that

$$\|s_N - h\|_{L^1} < \varepsilon.$$

Now,

$$\|F - h\|_{L^1} \leq \|F - s_N\|_{L^1} + \|s_N - h\|_{L^1} < 2\varepsilon.$$

Thus, for a given ε, for $\mid f \mid$ and f^-, there are continuous functions, h and g, respectively, such that

$$\||f| - h\|_{L^1} < \varepsilon \text{ and } \|2f^- - g\|_{L^1} < \varepsilon.$$

Therefore, for $H = g - h$,

$$\begin{aligned} \|f - H\|_{L^1} &= \||f| - 2f^- + g - h\|_{L^1} \\ &\leq \||f| - h\|_{L^1} + \|2f - g\|_{L^1} < 2\varepsilon. \end{aligned}$$ □

3.13 Let A be a subset of $[a, b]$ whose closure is negligible w.r.t. Borel - Lebesgue measure λ. Is the function $g = \mathbf{1}_A$ Riemann integrable on $[a, b]$? If it is Riemann integrable, then calculate $(\mathrm{R})\int_a^b g(x)dx$.

Solution. (i) Consider the function $f = \mathbf{1}_{\overline{A}}$. Obviously, $f = 0$ on $B = \overline{A}^c$ which is an open set. Consequently, f is continuous on B. (Note that this is not the case with the Dirichlet function $\mathbf{1}_{\mathbb{Q}}$ which is also constant on a subset that is not open; this is essential.) Indeed, because f is a constant function on B, say $f = c$, and if $x \in B$ is a point and O_c is any open neighborhood of $f(x) = c$, obviously,

$$f^*(O_c) = f^*(\{c\}) = B.$$

Now, because B is an open set and $x \in B$, it shows that x is an interior point of $B = f^*(O_c)$. Thus, f is continuous at any point $x \in B$.

(ii) Because $\overline{A}$ is closed, it is a Borel subset of $[a, b]$ and therefore, f is a Borel function on $[a, b]$. According to (i), f is continuous on $[a, b]$ λ-a.e. Thus, by Lebesgue Theorem 3.5, Chapter 6, f is Riemann integrable. By Theorem 3.2, f is also Lebesgue integrable and $(\mathrm{L})\int f = (\mathrm{R})\int f$.

(iii) Because $f \in [\mathbf{0}]_\lambda$,

$$\int_a^b f d\lambda = 0 = \int_a^b f(x)dx.$$

(iv) Now, we return to function $g = \mathbf{1}_A$. Obviously, $g = f$ λ-a.e. because $A \subseteq \overline{A}$ and $0 \leq g \leq f$. By Problem 5.10, Chapter 5, then g must also be λ-measurable and thus Lebesgue integrable. From (iii), in particular, it follows that $\int_a^b g\, d\lambda = 0$.

(v) Finally, is g Riemann integrable? Because $g = f$ on B and f is continuous on B, it follows that g is also continuous on B and thus on $[a, b]$ λ-a.e. Thus, g must be Riemann integrable and from (iii) and Theorem 3.2, thus $\int_a^b g(x)dx = 0$. □

4.1 Show that $(f \circ F)^+ = f^+ \circ F$ and $(f \circ F)^- = f^- \circ F$.

Solution. (i) Suppose $[\Omega_0, \Omega, F]$ is surjective. Let

$$P = \{f > 0\} = f^*((0, \infty])$$

and

$$N = \{f \leq 0\}.$$

Then,

$$f_*^+(P) = f_*(P) \subseteq (0, \infty]$$

and

$$f_*^+(N) = \{0\}.$$

Now, let $P_0 = F^*(P)$ and $N_0 = F^*(N)$. Obviously, the latter two sets form a partition of Ω_0 and in as much as F was assumed to be surjective,

$$F_*(P_0) = P \text{ and } F_*(N_0) = N.$$

Thus,

$$f_*^+ \circ F_*(P_0) = f_* \circ F_*(P_0) = f_*(P) \subseteq (0, \infty],$$

and

$$f_*^+ \circ F_*(N_0) = \{0\}.$$

On the other hand, $(f \circ F)^+(\omega_0) = 0$ if and only if $(f \circ F)(\omega_0) \leq 0$. The latter, in turn, is equivalent to $(f^+ \circ F)(\omega_0) = 0$ and this holds if and only if $\omega_0 \in N_0$. The rest is obvious and thus $(f \circ F)^+ = f^+ \circ F$.

The case $(f \circ F)^- = f^- \circ F$ is proven similarly.

If F is not surjective, $F_*(P_0) \subseteq P$ and $F_*(N_0) \subseteq N$ (possibly each being a proper subset). Either way, the equation $(f \circ F)^+ = f^+ \circ F$ still holds and so does $(f \circ F)^- = f^- \circ F$. For instance, F can be made surjective when taking its codomain for a range and considering Res$f$ on F's codomain instead of f. □

4.3 Let $f \in \mathcal{C}^{-1}(\mathbb{R}^n, \mathcal{B}^n)$, $\alpha \in \mathbb{R}$, $b \in \mathbb{R}^n$, and $B \in \mathcal{B}^n$. Show that

$$\int_B f(\alpha x + b)\lambda(dx) = \frac{1}{|\alpha|^n} \int_A f(x)\lambda(dx),$$

where $A = \alpha B + b$.

Solution. Let $L(x) = \alpha x + b$. Then,

$$I = \int_B f(\alpha x + b)\lambda(dx) = \int_B f \circ L(x)\lambda(dx),$$

Let $A = L_*(B)$. Representing the Borel set B as

$$B = L^* \circ L_*(B) = L^*(A) = \tfrac{1}{\alpha}(A - b),$$

we have

$$I = \int\limits_{L^*(A)} f \circ L(x)\lambda(dx)$$

[by (4.1a)]

$$= \int\limits_A f(x)\lambda L^*(dx).$$

By Proposition 4.3, Chapter 5, $\lambda L^* = \frac{1}{|\alpha|^n}\,\lambda$, implying that

$$I = \tfrac{1}{|\alpha|^n} \int\limits_A f(x)\lambda(dx), \tag{P4.3}$$

where $A = \alpha B + b$. Equation (P4.3) is due to Problem 1.18, that is, due to $\int f d(c\mu) = c\int f d\mu$, where $c > 0$. □

5.2 Let μ and ν be measures on (Ω, Σ) such that $\nu \leq \mu$ and suppose μ is σ-finite. Show that there exists a Radon - Nikodym density $f \in \frac{d\nu}{d\mu}$, such that $0 \leq f \leq 1$.

Solution. Because $\nu \leq \mu \;\Rightarrow\; \nu \ll \mu$. Thus by the Radon - Nikodym theorem, there is a Radon - Nikodym density $f \in \frac{d\nu}{d\mu} \geq 0$. Let

$$A : = \{f > 1\} = f^*((0, \infty)).$$

We show that A is the empty set. Let

$$A_n : = f^*\big([1 + \tfrac{1}{n}, \infty)\big).$$

Then, $\{A_n\} \uparrow$ and

$$\bigcup_{n=1}^{\infty} A_n = A.$$

Now,

$$f\mathbf{1}_{A_n} \geq \mathbf{1}_{A_n}\big(1 + \tfrac{1}{n}\big).$$

Furthermore,

$$\nu(A_n) = \int f \mathbf{1}_{A_n} d\mu \geq \left(1 + \tfrac{1}{n}\right)\mu(A_n)$$

and thus, $\nu(A_n) > \mu(A_n)$ for all $n = 1, 2, \ldots$, which of course contradicts the assumption that $\nu \leq \mu$. Therefore, each A_n must be the empty set. It obviously cannot be a μ-null set, nor can it be a ν-null set. Consequently, $A = \emptyset$. □

5.4. Let π_α be a Poisson measure on $(\mathbb{R}, \mathcal{B})$. Investigate whether π_α is absolutely continuous or singular relative to λ.

Solution. Let $B = \{0, 1, \ldots\}$. Then B is a Borel λ-null set. However, $\pi_\alpha(B) = 1$. Hence, π_α is not absolutely continuous relative to λ. Also, $\lambda(B) = \pi_\alpha(B^c) = 0 \;\Rightarrow\; \lambda \perp \pi_\alpha$. □

5.6 Let $\rho_1 \ll \mu$ and $\rho_2 \perp \mu$. Show that $\rho_1 \perp \rho_2$.

Solution. From $\rho_2 \perp \mu$ it follows that there is an A such that

$$\rho_2(A) = \mu(A^c) = 0.$$

On the other hand, $\rho_1 \ll \mu$ implies that $\rho_1(A^c) = 0$. Therefore, $\rho_1 \perp \rho_2$. □

5.8 Let μ and ν be σ-finite measures on (Ω, Σ). Show that μ and ν possess densities f and g, respectively, relative to $\rho = \mu + \nu$.

Solution. Clearly, $\rho(A) = 0 \Rightarrow \mu(A) = 0$ and $\nu(A) = 0$. Therefore, $\mu \ll \rho$ and $\nu \ll \rho$. Now, because either μ and ν is σ-finite, so is ρ. Thus, by Radon - Nikodym theorem 5.7, there are nonnegative Radon - Nikodym densities, say f and g from the unique equivalent classes $\frac{d\mu}{d\rho}$ and $\frac{d\nu}{d\rho}$, respectively. □

5.10. Let $f \in \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+))$, μ be a positive finite measure, $\Sigma_0 \subseteq \Sigma$ be a sub-σ-algebra, and $\mu_0 = \text{Res}_{\Sigma_0}\mu$. Suppose $\Sigma_0 = \{\Omega, \emptyset, A, B = A^c\}$ where A is a measurable set. Define

$$g = \tfrac{1}{\mu_0(A)} \int\limits_A f d\mu_0 \mathbf{1}_A + \tfrac{1}{\mu_0(B)} \int\limits_B f d\mu_0 \mathbf{1}_B.$$

Show that g is a Radon - Nikodym density of $\nu = \int f d\mu$ with respect to μ_0.

Solution. Clearly, g is Σ_0-measurable and μ-integrable. Now,

$$g\mathbf{1}_A = \tfrac{1}{\mu_0(A)} \int\limits_A f d\mu_0 \mathbf{1}_A \text{ and } g\mathbf{1}_B = \tfrac{1}{\mu_0(B)} \int\limits_B f d\mu_0 \mathbf{1}_B.$$

Therefore,

$$\int\limits_A g d\mu_0 = \int \mathbf{1}_A g d\mu_0 = \tfrac{1}{\mu_0(A)} \int\limits_A f d\mu_0 \, \mu_0(A) = \int\limits_A f d\mu_0$$

and similarly

$$\int\limits_B g d\mu_0 = \int\limits_B f d\mu_0.$$

We also have

$$\begin{aligned} \int\limits_\Omega g d\mu_0 &= \tfrac{1}{\mu_0(A)} \int\limits_A f d\mu_0 \, \mu_0(A) \\ &\quad + \tfrac{1}{\mu_0(B)} \int\limits_B f d\mu_0 \, \mu_0(B) \\ &= \int\limits_\Omega f d\mu_0. \end{aligned}$$

Therefore, $\int g d\mu_0 = \int f d\mu_0$ on Σ_0. □

5.12 Prove that $\mu \perp \rho$ if and only if $\mu \wedge \rho = \theta$.

Solution. (i) Let $\mu \perp \rho$. Then, there is a set A such that

$$\mu(A) = \rho(A^c) = 0.$$

Thus,

$$0 \leq (\mu \wedge \rho)(\Omega) = (\mu \wedge \rho)(A) + (\mu \wedge \rho)(A^c)$$

$$\leq \mu(A) + \rho(A^c) = 0.$$

Therefore, $\mu \wedge \rho = \theta$.

(ii) Let $\mu \wedge \rho = \theta$ hold. If $\mu \wedge \rho = \theta$, then given an $A \in \Sigma$, for each $\varepsilon > 0$, $\exists\, B$ such that $\mu(A) + \rho(A \backslash B) < \varepsilon$. By Lemma 1.11 and Theorem 1.13, Chapter 5,

$$0 = (\mu \wedge \rho)(A)$$

$$= \inf\{\mu(B) + \rho(A \backslash B) : B \subseteq A\}$$

implies that for each $\varepsilon > 0$, $\exists\, B \subseteq A$ such that $\mu(A) + \rho(A \backslash B) < \varepsilon$. In particular, for $A = \Omega$,

$$0 = (\mu \wedge \rho)(\Omega)$$

$\Rightarrow \forall\, \varepsilon > 0\ \exists\, E$ such that

$$\mu(E) + \rho(E^c) < \varepsilon;$$

that is, for each n, there is an E_n, such that

$$\mu(E_n) + \rho(E_n^c) < 2^{-n}.$$

Let $A_n = \bigcup_{i=n}^{\infty} E_i$ and $A = \bigcap_{n=1}^{\infty} A_n$. Then,

$$\mu(A) \leq \mu(A_n) = \mu\left(\bigcup_{i=n}^{\infty} E_i\right)$$

$$\leq \textstyle\sum_{i=n}^{\infty} \mu(E_i) \leq \sum_{i=n}^{\infty} 2^{-i} = 2^{1-n}.$$

Thus, if $n \to \infty$, $\mu(A) = 0$. On the other hand,

$$\rho(A_n^c) = \rho\left(\bigcap_{i=n}^{\infty} E_i^c\right) \leq \rho(E_i^c) \leq \tfrac{1}{2^i}$$

$\forall i \geq n$ independent of n. Hence,

$$\rho(A_n^c) = 0$$

implying that

$$\rho\left(\bigcup_{n=1}^{\infty} A_n^c\right) \leq \textstyle\sum_{n=1}^{\infty} \rho(A_n^c) = \rho(A^c) = 0. \qquad \square$$

6.2 Let ν be a signed measure on a measurable space (Ω, Σ) and let A be a ν-finite measurable set. Show that if $A = A_1 + A_2$ is any decomposition of A, then either A_1 and A_2 is ν-finite. Furthermore, show that if $\nu(A) = \infty$, then at least one of the subsets A_1 or A_2 must be positively infinite. If $A = \Sigma_{n=1}^{\infty} A_n$ is a countable measurable decomposition of a finite set A, then all sets A_n's are finite. In particular, it follows that any subset of a finite set must be finite.

Solution. Indeed, by additivity of ν, $\nu(A) = \nu(A_1) + \nu(A_2)$, and thus it is impossible that at least one of the terms on the right is infinite. Besides, by our earlier convention, we cannot operate with $+\infty - \infty$, which would make the additivity property of ν inconsistent. Now, if $\nu(A) = \infty$, then again due to the consistency of ν, at least one of the subsets A_1 or A_2 must be positively infinite.

If A is finite and at least one set , say A_k, is positively infinite, then σ-additivity does not hold.

Any subset S of a finite set A is finite because $A = S + (A \backslash S)$ is a decomposition of A which we argued above. $\square$

6.6 Prove Proposition 6.6. *If A and B are measurable sets such that B is ν-finite and $A \subseteq B$, then A is also ν-finite and $\nu(B\backslash A) = \nu(B) - \nu(A)$.*

Solution. From $B = (B\backslash A) + A$ and additivity of ν we have $\nu(B) = \nu(B\backslash A) + \nu(A)$. By Problem 6.2, either term on the right must be finite. The latter allows us to subtract $\nu(A)$ from either side of the equation to yield the statement. □

6.8 Prove Proposition 6.8. *Let (Ω, Σ, ν) be a signed measure space and $\{A_n\}$ be a monotone nonincreasing sequence of measurable sets such that for some $n \in \mathbb{N}$, $\nu(A_n)$ is finite. Then the signed measure ν is continuous from above on $\{A_n\}$.*

Solution. (a) Assume that ν is a signed measure on Σ and that $\{A_n\}\downarrow$ with $|\nu(A_1)| < \infty$. Denote $A\colon = \bigcap_{n=1}^{\infty} A_n$. Then, by Problems 6.2 and 6.6, $|\nu(A_n)| < \infty$ for all $n > 1$ and thus $|\nu(A)| < \infty$. Because $\{A_n\}$ is monotone nonincreasing, $\{A_1\backslash A_n\}\uparrow$. It can easily be shown that

$$\bigcup_{n=1}^{\infty} (A_1\backslash A_n) = A_1\backslash\left(\bigcap_{n=1}^{\infty} A_n\right). \tag{1}$$

Now we apply Problem 6.7 to the sequence $\{A_1\backslash A_n\}$ and equation (1) to arrive at

$$\lim_{n\to\infty} \nu(A_1\backslash A_n) = \nu\left(\bigcup_{n=1}^{\infty} (A_1\backslash A_n)\right) = \nu(A_1\backslash A). \tag{2}$$

By our assumption ν is finite on $\{A_n\}$. Therefore, by Proposition 6.6, equation (2) reduces to

$$\nu(A_1) - \lim_{n\to\infty} \nu(A_n) = \nu(A_1) - \nu(A)$$

and thereby yields assertion (i).

(b) Now suppose $\nu(A_n)$ is finite for some n. Then, by Problems 6.2 and 6.6, A_k are also finite for all $k > n$ and thus we can apply the arguments of part (a) to the tail of the sequence $\{A_k : k \geq n\}$ and have (considering that $A = \bigcap_{k=n}^{\infty} A_k$):

$$\nu(A) = \lim_{k \to \infty} \nu(A_k). \qquad \square$$

6.10 Let $\mathfrak{P}$ and $\mathfrak{N}$ be the families of positive and negative sets, respectively. Show that $\mathfrak{P}$ and $\mathfrak{N}$ contain the empty set and they are closed with respect to differences and at most countable unions and intersections.

Solution. We render the proof for family $\mathfrak{P}$. It is obvious that the empty set is a trivial example of a positive set.

(i) Let $P, Q \in \mathfrak{P}$. Then, from the representation

$$(P \backslash Q) \cap A = P \cap (A \backslash Q)$$

[Problem 1.2e), Chapter 1] it follows that $P \backslash Q \in \mathfrak{P}$.

(ii) Let $\{P_n\} \subseteq \mathfrak{P}$ be a disjoint sequence. Then, by σ-additivity of ν,

$$\nu((\textstyle\sum_{n=1}^{\infty} P_n) \cap A) = \sum_{n=1}^{\infty} \nu(P_n \cap A)$$

is the sum of measures of sets $\nu(P_n \cap A) \geq 0,$ so that the sum

$$\nu((\textstyle\sum_{n=1}^{\infty} P_n) \cap A) \in [0, \infty].$$

(iii) If $\{P_n\}$ is an arbitrary sequence of positive sets, then we can set

$$P_1^* = P_1 \text{ and } P_n^* = P_n \backslash \big(\textstyle\sum_{i=1}^{n-1} P_i^*\big),\ n \geq 2,$$

so that $\{P_n^*\} \subseteq \mathfrak{P}$ [by parts (i) and (ii)] and by part (ii)

$$\bigcup_{i=1}^{\infty} P_i = \sum_{i=1}^{\infty} P_i^* \in \mathfrak{P}.$$

(iv) Let $\{P_n\} \subseteq \mathfrak{P}$ and denote $P_0 = \bigcap_{i=1}^{\infty} P_i$. Then,

$$P_0 \subseteq P_1 \Rightarrow \nu(P_0) \geq 0.$$

Furthermore, for any $A \in \Sigma$, $P_0 \cap A \subseteq P_1$ implying that $\nu(P_0 \cap A) \geq 0$. Therefore, $P_0 \in \mathfrak{P}$.

The arguments about the family of negative sets are identical. □

6.11 Let $\mathfrak{P}$ and $\mathfrak{N}$ be the families of positive and negative sets, respectively. Show that ν is monotone on either family.

Solution. Let $P, Q \in \mathfrak{P}$ such that $P \subseteq Q$. Then, from the decomposition $Q = P + (Q \backslash P)$ and the fact that $Q \backslash P \in \mathfrak{P}$ (by Problem 6.10) it follows that $\nu(P) \leq \nu(Q)$.

Now, let $M, N \in \mathfrak{N}$ such that $M \subseteq N$. Then, $N = M + (N \backslash M)$ is a decomposition of N into two negative subsets. Therefore, $\nu(N) \leq \nu(M)$. So, here we have the monotone decreasing property: with $M \subseteq N \Rightarrow \nu(M) \geq \nu(N)$. □

6.13 Show that all Hahn decompositions form a unique equivalence class.

Solution. Suppose (P', N') is another Hahn decomposition. Because $P \backslash P' \subseteq P$, by Remark 6.5$(iii)$, it is a positive set. But because $P \backslash P'$ is also representable in the form $P \cap (P')^c = P \cap N'$, it is a subset of N' and by the same Remark 6.5(iii), must be negative. Finally, by Remark 6.5(iii), (b), $P \backslash P'$ is a null set. Consequently, $P' \backslash P$ is a null set and thus $P \Delta P'$ is a null set. The rest is similar. □

6.15 Show that any signed Borel - Lebesgue - Stieltjes measure on $(\mathbb{R}^n, \mathcal{B})$ is σ-finite.

Solution. Recall that $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \nu)$ is a Borel - Lebesgue - Stieltjes measure space if ν is a Borel measure and it is finite on any d_e-bounded Borel sets. Obviously,

$$\nu\left(\prod_{i=1}^{n}(-a_i, a_i)\right) < \infty,$$

where $a_i = k$, and that

$$\left\{\prod_{i=1}^{n}(-a_i, a_i)\right\} \uparrow \mathbb{R}^n. \qquad \square$$

6.17 Let $\nu \in \mathfrak{S}_*(\Omega, \Sigma)$; that is, ν is a finite signed measure. Is it true that ν is bounded?

Solution.

$$|\nu(A)| \leq |\nu|(A)$$

$$= \nu^+(A) + \nu^-(A) \leq \nu^+(\Omega) + \nu^-(\Omega)$$

$$= \nu(P) - \nu(N) = \|\nu\| = M,$$

given $(P, N) \in [H^+, H^-]_\nu$. Because ν is finite, $\nu(P) < \infty$ and $-\nu(N) < \infty$. Thus, M is a nonnegative real number and we have that for each $A \in \Sigma$, $|\nu(A)| \leq M$. In particular, it follows that $\|\nu\| < \infty$. $\square$

6.19 Prove Lemma 6.25. *Let $\nu_1, \nu_2 \in \mathfrak{S}_*(\Omega, \Sigma)$. Then the set function*

$$\rho(A) = \sup\{\nu_1(E) + \nu_2(A \backslash E) : E \in \Sigma \cap A\}$$

is a finite signed measure on Σ.

Solution.

(i) Obviously, ρ is a set function defined on Σ that equals zero at the empty set. Next we show that ρ is finite. By definition,

$$\rho(\Omega) = \sup\{\nu_1(E) + \nu_2(E^c) : E \in \Sigma\}.$$

Then we have

$$\begin{aligned} &|\nu_1(E) + \nu_2(E^c)| \\ &\leq |\nu_1(E)| + |\nu_2(E^c)| \leq |\nu_1|(\Omega) + |\nu_2|(\Omega) \\ &= \|\nu_1\| + \|\nu_2\| < \infty. \end{aligned}$$

Therefore,

$$|\rho(\Omega)| \leq \|\nu_1\| + \|\nu_2\| < \infty.$$

(ii) Now we need to prove that ρ is σ-additive. Let $\{A_n\} \subseteq \Sigma$ be a sequence of disjoint sets with $A := \sum_{n=1}^{\infty} A_n$. Let $\varepsilon > 0$ be chosen. From the definition of ρ, given A_n, and an $\varepsilon_n > 0$, there is a subset $E_n \subseteq A_n$ such that

$$\rho(A_n) < \nu_1(E_n) + \nu_2(A_n \backslash E_n) + \varepsilon_n,\ n = 1, 2, \ldots$$

We choose ε_n's such that $\sum_{n=1}^{\infty} \varepsilon_n \leq \varepsilon$, which we specify beforehand. Denote $E = \sum_{n=1}^{\infty} E_n$. Obviously, E_n's are disjoint, $E \subseteq A$, and

$$\begin{aligned} A \backslash E &= \textstyle\sum_{n=1}^{\infty} (A_n \backslash \sum_{k=1}^{\infty} E_k) \\ &= \textstyle\sum_{n=1}^{\infty} (A_n \backslash E_n), \end{aligned}$$

because the rest of the E_k's are disjoint with A_n. Therefore, with

$$\begin{aligned} &\nu_1(E) + \nu_2(A \backslash E) \\ &\quad = \textstyle\sum_{n=1}^{\infty} \nu_1(E_n) + \sum_{n=1}^{\infty} \nu_2(A_n \backslash E_n) \end{aligned}$$

we have:

$$\textstyle\sum_{n=1}^{\infty} \rho(A_n) \leq \sum_{n=1}^{\infty} \nu_1(E_n)$$

$$
\begin{aligned}
&+ \textstyle\sum_{n=1}^{\infty} \nu_2(A_n \backslash E_n) + \sum_{n=1}^{\infty} \varepsilon_n \\
&\leq \nu_1(E) + \nu_2(A \backslash E) + \varepsilon \\
&\leq \rho(A) + \varepsilon.
\end{aligned}
$$

Because this inequality holds for each $\varepsilon > 0$, we arrive at the desired inequality

$$\textstyle\sum_{n=1}^{\infty} \rho(A_n) \leq \rho(A).$$

In particular, we notice that $\sum_{n=1}^{\infty} \rho(A_n) < \infty$ that holds true for any disjoint sequence of sets $\{A_n\}$.

(iii) Now, we establish the inverse inequality. We thus need to show that ρ is σ-subadditive on $\{A_n\}$, i.e.

$$\rho(A) \leq \textstyle\sum_{n=1}^{\infty} \rho(A_n). \tag{P6.19}$$

Let E be a measurable subset of A. Then,

$$
\begin{aligned}
&\nu_1(E) + \nu_2(A \backslash E) \\
&= \nu_1(\textstyle\sum_{n=1}^{\infty}(A_n \cap E)) + \nu_2(\sum_{n=1}^{\infty}(A_n \cap E^c)) \\
&= \textstyle\sum_{n=1}^{\infty} \nu_1(A_n \cap E) + \nu_2(A_n \backslash (A_n \cap E)) \\
&\leq \textstyle\sum_{n=1}^{\infty} \rho(A_n)
\end{aligned}
$$

which yields (P6.19). The series $\sum_{n=1}^{\infty} \rho(A_n) < \infty$ as noticed from part (ii). On the other hand, it also follows that

$$-\infty < \textstyle\sum_{n=1}^{\infty} \rho(A_n).$$

So, we are done with the lemma. □

6.21 Let $\nu, \rho \in \mathfrak{S}_*(\Omega, \Sigma)$. Show that $\nu + \rho \in \mathfrak{S}_*(\Omega, \Sigma)$. Is it true that

$$\nu + \rho = (\nu + \rho)^+ - (\nu + \rho)^-$$

$$= \nu^+ + \rho^+ - \nu^- - \rho^-?$$

Solution. Obviously, $\nu + \rho$ is a finite signed measure through its representation $\nu + \rho = \nu^+ + \rho^+ - \nu^- - \rho^-$. Given a measurable set A, we define $(\nu + \rho)(A) = \nu(A) + \rho(A)$. However,

$$\nu^+ + \ \rho^+ - \nu^- - \rho^-$$

need not be a Hahn decomposition of $\nu + \rho$. This is because the Jordan variations ν^+, ν^- and ρ^+, ρ^- relate to possibly "entirely" different Hahn classes $[H^+, H^-]_\nu$ and $[H^+, H^-]_\rho$. Or equivalently, $\nu^+ + \rho^+$ need not be orthogonal to $\nu^- + \rho^-$, even though $\nu^+ \perp \nu^-$ and $\rho^+ \perp \rho^-$ hold. □

6.23 Show that $\mathfrak{S}_*(\Omega, \Sigma)$ is a vector space over the field $\mathbb{R}$.

Solution. The statement is in part due to Problems 6.21 and 6.23. In addition, the zero measure θ must be included. □

7.1 Show that the total variation of a complex measure ξ has the property, $|\xi| \leq \nu_1^+ + \nu_1^- + \nu_2^+ + \nu^-$ and that in general, $|\xi| < \nu_1^+ + \nu_1^- + \nu_2^+ + \nu^-$. Thus, it disagrees with the associated property for a signed measure.

Solution.(i) First, from (7.3a) we have that

$$|\xi|(A) \leq |\nu_1|(A) + |\nu_2|(A) = (\nu_1^+ + \nu_1^- + \nu_2^+ + \nu_2^-)(A).$$

(ii) The reason why $|\xi| < \nu_1^+ + \nu_1^- + \nu_2^+ + \nu^-$ is purely algebraic. For a set E,

$$|\xi(E)| = \sqrt{(\nu_1^+(E) - \nu_1^-(E))^2 + (\nu_2^+(E) - \nu^-(E))^2},$$

in general

$$< \Big[(\nu_1^+(E) - \nu_1^-(E))^2 + 2|\nu_1^+(E) - \nu_1^-(E)|$$

$$\times |\nu_2^+(E) - \nu_2^-(E)| + (\nu_2^+(E) - \nu^-(E))^2\Big]^{1/2}.$$

Consider the complex measure space $([0,4], \mathcal{B}([0,4]), \xi)$, where

$$\xi = \lambda - 2\varepsilon_0 + i(\varepsilon_1 - \varepsilon_2).$$

It is easy to verify that ξ is a complex measure as it has two signed measure components. Consider the partition

$$E_1 = \{0\},\ E_2 = (0,2),\ E_3 = [2,4]$$

of $[0,4]$ and calculate $\sum_{i=1}^3 |\xi(E_i)| = \sqrt{2} + 2\sqrt{5}$. It can be readily shown that

$$|\xi|([0,4]) = \textstyle\sum_{i=1}^3 |\xi(E_i)|,$$

whereas the sum of all Jordan variations of set $[0,4]$ is

$$(\lambda + 2\varepsilon_0 + \varepsilon_1 + \varepsilon_2)([0,4]) = 8. \qquad \square$$

7.3 Prove Proposition 7.3. *Any complex measure ξ is continuous from below.*

Solution. Let ξ be a complex measure on Σ and $\{A_n\}$ a monotone nondecreasing sequence in Σ such that $A := \bigcup_{n=1}^{\infty} A_n$. Denote

$$B_n = A_n \backslash A_{n-1}\,,\ n = 1, 2, \ldots,\ A_0 = \emptyset\,.$$

Then

$$A_n = \textstyle\sum_{k=1}^n B_k$$

and $\{B_n\}$ forms a pairwise disjoint sequence from Σ with

$$\textstyle\sum_{n=1}^{\infty} B_n = A.$$

Because $|\xi(A_n)| < \infty$, $\xi(B_n)$ is well defined and, therefore, due to σ-additivity of ξ and because $|\xi(A)| < \infty$,

$$\begin{aligned}\lim_{n\to\infty}\xi(A_n) &= \lim_{n\to\infty}\textstyle\sum_{k=1}^{n}\xi(B_k)\\ &= \textstyle\sum_{k=1}^{\infty}\xi(B_k)\\ &= \xi\left(\textstyle\sum_{k=1}^{\infty}B_k\right) = \xi(A),\end{aligned}$$

which yields continuity from below. □

7.5 Prove Proposition 7.5. *Let ξ be a finitely additive set function on the measurable space (Ω, Σ) such that $\xi(\emptyset) = 0$. ξ is continuous from below if and only if ξ is σ-additive.*

Solution. (i) Suppose ξ is continuous from below and let $\{B_n\}$ be a sequence of pairwise disjoint measurable sets from Σ. Let

$$A_n = \textstyle\sum_{i=1}^{n} B_i.$$

Then, $\{A_n\}\uparrow \subseteq \Sigma$ and $A = \bigcup_{n=1}^{\infty} A_n$ equals $\sum_{n=1}^{\infty} B_n$. Due to continuity from below of ξ and by its finite additivity we have σ-additivity:

$$\begin{aligned}\xi(A) &= \xi\left(\textstyle\sum_{n=1}^{\infty}B_n\right) = \lim_{n\to\infty}\xi(A_n)\\ &= \lim_{n\to\infty}\textstyle\sum_{i=1}^{n}\xi(B_i) = \sum_{i=1}^{\infty}\xi(B_i). \qquad (1)\end{aligned}$$

(ii) Let $\{B_n\}$ be a sequence of measurable pairwise disjoint sets. We form the sequence $\{A_n\}\uparrow$ as in part (i). Then $\{A\backslash A_n\}\downarrow \emptyset$ and continuity from above yields $\emptyset$-continuity and the latter in turn yields

$$\lim_{n\to\infty}\xi(A\backslash A_n) = \xi(\emptyset) = 0.$$

Because the ξ is finite on Σ, by Problem 7.2, the last expression leads to

$$\lim_{n \to \infty} \xi(A_n) = \xi(A).$$

Applying equations (1) we arrive at σ-additivity of ξ.

(iii) Suppose ξ is Ø-continuous from above. Let $\{A_n\}$ be a monotone nonincreasing sequence of measurable sets and let $A = \bigcap_{n=1}^{\infty} A_n$. Let $B_n := A_n \backslash A$. Then, obviously, $\bigcap_{n=1}^{\infty} B_n = \emptyset$. Now we have

$$\xi(A_n) = \xi(B_n + A) = \xi(B_n) + \xi(A).$$

By the above assumption,

$$\lim_{n\to\infty}\xi(A_n) = \lim_{n\to\infty}\xi(B_n) + \xi(A) = 0 + \xi(A).$$

Thus, ξ is continuous from above. By part (ii), ξ is therefore σ-additive. □

7.7 Any complex measure $\xi \in \mathcal{C}(\Omega, \Sigma)$ is finite. It means that for any measurable set A, $|\xi(A)| < \infty$. (a) Does it mean that ξ is bounded? (b) Let $\xi = \nu_1 + i\nu_2$, where $\nu_1, \nu_2 \in \mathfrak{S}_*(\Omega, \Sigma)$. Show that for any $A \in \Sigma$, $|\xi(A)| \leq \|\nu_1\| + \|\nu_2\|$.

Solution. Given $A \in \Sigma$ and $\xi = \nu_1 + i\nu_2$ we have

$$|\xi(A)|^2 = |\nu_1(A)|^2 + |\nu_2(A)|^2 \leq (|\nu_1(A)| + |\nu_2(A)|)^2$$

and thus

$$|\xi(A)| \leq |\nu_1(A)| + |\nu_2(A)| \tag{P7.7}$$

$$\leq |\nu_1|(\Omega) + |\nu_2|(\Omega) = \|\nu_1\| + \|\nu_2\|.$$

The latter is a real number, according to Problem 6.17. □

7.9 Let $\xi, \xi_1, \xi_2 \in \mathcal{C}(\Omega, \Sigma)$ such that for each $A \in \Sigma$, $|\xi(A)| \leq |\xi_1(A)| + |\xi_2(A)|$. Prove that $\|\xi\| \leq \|\xi_1\| + \|\xi_2\|$.

Solution. Let $\mathcal{S}(A) \subseteq \Sigma$ be the family of all finite subpartitions of A and let $\{E_1, \ldots, E_n\} \subseteq \mathcal{S}(A)$. Then,

$$\begin{aligned}\textstyle\sum_{i=1}^n |\xi(E_i)| &\leq \textstyle\sum_{i=1}^n |\xi_1(E_i)| \\ &\quad + \textstyle\sum_{i=1}^n |\xi_2(E_i)| \\ &\leq |\xi_1|(A) + |\xi_2|(A).\end{aligned}$$

Therefore,

$$|\xi|(A) \leq |\xi_1|(A) + |\xi_2|(A)$$

implying the desired inequality $\|\xi\| \leq \|\xi_1\| + \|\xi_2\|$. □

7.11 Let $\xi = \nu + i\rho$ be a complex measure on a measurable space (Ω, Σ). Prove the following inequalities for total variations and norms.

$$|\xi| \geq \tfrac{1}{\sqrt{2}}(|\nu| + |\rho|) \text{ and } \|\xi\| \geq \tfrac{1}{\sqrt{2}}(\|\nu\| + \|\rho\|).$$

Solution. We first notice that if $\xi = \nu + i\rho$ is a complex measure, then for each measurable set A,

$$\sqrt{2}|\xi(A)| \geq |\nu(A)| + |\rho(A)|. \tag{P7.11}$$

Furthermore, let $\mathcal{S}(A)$ be a set of all finite measurable subpartitions of set A and let $\{E_1, \ldots, E_n\} \subseteq \mathcal{S}(A)$. Then, by (P7.11),

$$\begin{aligned}&\textstyle\sum_{i=1}^n |\nu(E_i)| + \sum_{i=1}^n |\rho(E_i)| \\ &\leq \sqrt{2}\textstyle\sum_{i=1}^n |\xi(E_i)| \leq \sqrt{2}|\xi|(A).\end{aligned}$$

Therefore, we conclude that for any measurable set A,

$$|\nu|(A) + |\rho|(A) \leq \sqrt{2}|\xi|(A).$$

The similar normed inequality follows immediately when setting $A = \Omega$. □

7.13 Prove Theorem 7.8 that $(\mathcal{C}(\Omega, \Sigma), \|\cdot\|)$ is a Banach space.

Solution. Let $\{\xi_n = \nu_n + i\rho_n\} \subseteq \mathcal{C}(\Omega, \Sigma)$ be a Cauchy sequence. Then, by Problem 7.12, $\{\nu_n\}$ and $\{\rho_n\}$ are Cauchy sequences of signed measures. By Theorem 6.31, we have the following.

(i) $\{\nu_n\}$ and $\{\rho_n\}$ converge to signed measures ν and ρ uniformly setwise,

(ii) $\mathfrak{S}_*(\Omega, \Sigma)$ is complete and $\{\nu_n\}$ and $\{\rho_n\}$ converge to ν and ρ in the norm.

Therefore, we define $\xi = \nu + i\rho$. Obviously, ξ is a complex measure. Now, $\xi_n - \xi = \nu_n - \nu + i(\rho_n - \rho)$ and thus by Problem 7.8, for each $\varepsilon > 0$,

$$\|\xi_n - \xi\| \leq \|\nu_n - \nu\| + \|\rho_n - \rho\| < \varepsilon,\ n \geq N,$$

implying that $\mathcal{C}(\Omega, \Sigma)$ is complete in $\|\cdot\|$. □

7.15 Show that, given a signed measure ν, it holds true that $L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}}) = L^1(\Omega, \Sigma, |\nu|; \overline{\mathbb{R}})$.

Solution. By (7.9a) $f \in L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}})$ if and only if

$$f \in\ L^1(\Omega, \Sigma, \nu^+; \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma, \nu^-; \overline{\mathbb{R}}).$$

On the other hand, because $|\nu| = \nu^+ + \nu^-$, by Example 1.9 (iv),

$$\int |f| d|\nu| = \int |f| d(\nu^+ + \nu^-)$$

$$= \int |f| d\nu^+ + \int |f| d\nu^-.$$

Therefore, f or $|f| \in L^1(\Omega, \Sigma, |\nu|; \overline{\mathbb{R}})$ if and only if

$$f \in\ L^1(\Omega, \Sigma, \nu^+; \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma, \nu^-; \overline{\mathbb{R}}).$$

□

7.17 Let (Ω, Σ, μ) be a measure space, with μ being a positive measure. Prove that for each $f \in L^1(\Omega, \Sigma, \mu; \mathbb{C})$, $|\xi| = \int |f| d\mu$, where $\xi = \int f d\mu$.

Solution. (i) First we show that $|\xi| \leq \int |f| d\mu$. Let $A \in \Sigma$ and $\{E_1, \ldots, E_k\} \in \mathcal{S}(A)$, where the latter is the set of all measurable subpartitions of A. Then, by Problem 7.14,

$$\sum_{i=1}^k |\nu(E_i)| = \sum_{i=1}^k \left| \int_{E_i} f d\mu \right| \leq \int_A |f| d\mu.$$

(ii) We show the inverse inequality. Define

$$g = \frac{\overline{f}}{|f|} \quad \text{and} \quad \widetilde{g} = \frac{\overline{f}}{|f|} \mathbf{1}_{\{|f|>0\}}.$$

Then, because $\widetilde{g}$, as a complex-valued function, consists of four positive components (s^+, s^-, t^+, t^-), there are four monotone increasing sequences of simple functions that converge pointwise to (s^+, s^-, t^+, t^-). In other words, there is a sequence $\{g_n\}$ of "complex-valued simple functions" that converge pointwise to $\widetilde{g}$. We can say that a sequence in polar form

$$\left\{ g_n = \sum_{j=1}^{m_n} r_j^{(n)} e^{i\theta_j^{(n)}} \mathbf{1}_{A_j^{(n)}} \right\}$$

converges to $\widetilde{g}(\omega) = r(\omega) e^{i\theta(\omega)}$ pointwise. [Notice that $r(\omega) = \mathbf{1}_{\{|f(\omega)|>0\}}$.] Therefore, the sequence of functions $\{g_n f\}$ converges pointwise to the function $\widetilde{g} f = |f| \mathbf{1}_{\{|f|>0\}}$. Therefore, by the Lebesgue dominated convergence theorem (for complex-valued functions, as of Problem 7.16), the integral $\int g_n f d\mu$ converges to $\int \widetilde{g} f d\mu$ and the latter obviously equals

$$\int g f d\mu = \int |f| d\mu.$$

Notice that because $|\widetilde{g}| = 1$ on $\{|f| > 0\}$, the sequence $\{g_n\}$ must be bounded and so is $\{g_n f\}$ μ-a.s. on Ω.

Now, if g_n is in the form

$$g_n = \sum_{j=1}^{m_n} \gamma_j^{(n)} \mathbf{1}_{A_j^{(n)}} = \sum_{j=1}^{m_n} r_j^{(n)} e^{i\theta_j^{(n)}} \mathbf{1}_{A_j^{(n)}},$$

then we have

$$\begin{aligned}\left| \int_A g_n f d\mu \right| &= \left| \int_A \sum_{j=1}^{m_n} r_j^{(n)} e^{i\theta_j^{(n)}} \mathbf{1}_{A_j^{(n)}} f d\mu \right| \\ &\le \sum_{j=1}^{m_n} r_j^{(n)} \left| \xi\left(A \cap A_j^{(n)}\right) \right| \\ &\le r_n \sum_{j=1}^{m_n} |\xi| \left(A \cap A_j^{(n)}\right) \\ &= r_n |\xi|(A), \end{aligned} \tag{P7.17}$$

where

$$r_n = \max\left\{ r_j^{(n)} : j = 1, \dots, m_n \right\}.$$

Because $r_j^{(n)} \to 1$ for all j's [because $r = 1$ on the set $\{|f| > 0\}$; because we deal with an inequality, we can assume $r = 1$ everywhere on Ω], so does $r_n \to 1$. So, the right-hand side of (P7.17) converges to $|\xi|(A)$, whereas the left-hand side of (P7.17), by the Lebesgue dominated convergence theorem (as noticed) converges to $\int_A |f| d\mu$. This shows the validity of the inverse inequality and herewith the equation

$$|\xi|(A) = \int_A |f| d\mu, \quad \text{for all } A \in \Sigma. \qquad \square$$

7.19 Let (Ω, Σ, ν) be a signed measure space and $f \in \mathbb{L}(\Omega, \Sigma; \nu, \bar{\mathbb{R}}_+)$. Define $\rho = \int f d\nu^+ - \int f d\nu^-$, $\rho_1 = \int f d\nu^+$, and $\rho_2 = \int f d\nu^-$. Prove that ρ_1 and ρ_2 are the Jordan variations of ρ.

Solution. The set function ρ is a signed measure as long as at least one of the integrals ρ_1 or ρ_2 exists, which we assume to be the case. Now we need to check if $\rho_1 \perp \rho_2$.

(a) Let $g = \mathbf{1}_A$, where $A \in \Sigma$ is fixed. Define

$$\rho_1 = \int g d\nu^+ = \int \mathbf{1}_A d\nu^+$$

and

$$\rho_2 = \int g d\nu^- = \int \mathbf{1}_A d\nu^-.$$

Then, for some Hahn pair $(P, N) \in [H^+, H^-]_\nu$,

$$\rho_1(N) = \nu^+(N \cap A) = 0$$

and

$$\rho_2(P) = \nu^-(P \cap A) = 0,$$

because ν^+ and ν^- are both positive measures and thus are monotone. Therefore, $\rho_1 \perp \rho_2$ through (P, N).

(b) Let $g = \sum_{i=1}^m a_i \mathbf{1}_{A_i}$. Then,

$$\rho_1(\,\cdot\,) = \sum_{i=1}^m a_i \nu^+(\,\cdot\, \cap A_i),$$

$$\rho_2(\,\cdot\,) = \sum_{i=1}^m a_i \nu^-(\,\cdot\, \cap A_i).$$

Again take for some Hahn pair $(P, N) \in [H^+, H^-]_\nu$,

$$\rho_1(N) = \sum_{i=1}^m a_i \nu^+(N \cap A_i)$$

$$= \rho_2(P) = \sum_{i=1}^m a_i \nu^-(P \cap A_i) = 0.$$

Therefore, in this case $\rho_1 \perp \rho_2$.

(c) Let $\{g_n\} \subseteq \Psi_+$ be a monotone nondecreasing sequence convergent pointwise to an $f \in \mathcal{C}_+^{-1}$ previously selected. Then, with $\rho_1 = \int f d\nu^+$ and $\rho_2 = \int f d\nu^-$ we have

$$\rho_1(N) = \sup \int \mathbf{1}_N g_n d\nu^+$$

$$= \rho_2(P) = \sup \int \mathbf{1}_P g_n d\nu^- = 0$$

due to case (b) for each g_n. Thus, $\rho_1 \perp \rho_2$. Therefore, $\rho = \rho_1 - \rho_2$ is the Jordan decomposition relative to the class $[H^+, H^-]_\nu$, as per Pro-

position 6.19, and ρ_1 and ρ_2 are indeed the Jordan variations of the integral measure ρ. □

7.21 Let (Ω, Σ, ν) be a signed measure space and $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$. Prove that $f = 0$ ν-a.e. if and only if $\int\limits_A f d\nu = 0$ for all $A \in \Sigma$.

Solution. Let $\rho = \int f d\nu = \int f d\nu^+ - \int f d\nu^- = \rho_1 - \rho_2$, where ρ_1 and ρ_2 are signed measures, because f is an extended real-valued function. According to Problem 7.16 which also holds for signed measures, $|\rho_1| = \int |f| d\nu^+$ and $|\rho_2| = \int |f| d\nu^-$. Furthermore,

$$|\rho_1| = \int |f| d\nu^+ \quad \text{and} \quad |\rho_2| = \int |f| d\nu^-$$

are positive integral measures with respect to the Jordan variations ν^+ and ν^-. Because $|f| \in \mathcal{C}_+^{-1}$, by Problem 7.19, $|\rho_1|$ and $|\rho_2|$ are the Jordan variations of the integral signed measure

$$|\rho_1| - |\rho_2| = \int |f| d\nu^+ - \int |f| d\nu^-$$

and thus, the latter is $\pi := \int |f| d\nu$. Furthermore, as the Jordan variations of the signed measure π, $|\rho_1| \perp |\rho_2|$.

Now, if $\rho = \theta$, as by the assumption that $\rho(A) = 0$ for all $A \in \Sigma$, we conclude that $\rho_1 = \rho_2$. But then also $|\rho_1| = |\rho_2|$ and because they are orthogonal, as we argued in Problem 7.20, $|\rho_1| = |\rho_2| = \theta$. This in turn yields that $|f| = 0$ ν^+-a.e. and $|f| = 0$ ν^--a.e. Consequently, $|f| = 0$ ν-a.e. and so is $f = 0$ ν-a.e.

The converse of this statement is trivial. □

7.23 Prove the monotone convergence theorem for a monotone nondecreasing sequence of nonnegative measurable functions $\{f_n\}$ on a signed measure space pointwise convergent to $f \in \mathbb{L}(\Omega, \Sigma; \nu, \overline{\mathbb{R}}_+)$, namely,

$$\lim_{n\to\infty} \int f_n d\nu = \int f d\nu.$$

Solution. Because $\sup\int f_n d\nu^+$ exists and equals $\int f d\nu^+$ by the monotone convergence theorem with respect to positive measures and so is $\sup\int f_n d\nu^- = \int f d\nu^-$ and if each is finite, we have

$$\lim_{n\to\infty}\int f_n d\nu = \int f d\nu^+ - \int f d\nu^- = \int f d\nu.$$

In the event one of the limits (just one only) is infinity, $\sup\int f_n d\nu$ still makes sense and it will equal plus or minus infinity. □

7.25 Let (Ω, Σ) be a measurable space and ν and ρ be two signed measures on (Ω, Σ) such that $\nu + \rho$ is defined. Suppose a and b are two real constants. (i) Is the linear combination $a\nu + b\rho$ always defined? (ii) If $a\nu + b\rho$ is defined, show that it is a signed measure on (Ω, Σ) and prove that for any $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma; \rho, \overline{\mathbb{R}})$,

$$\int f d(a\nu + b\rho) = a\int f d\nu + b\int f d\rho.$$

Solution. (i) No, it is not always defined if $\nu + \rho$ is defined. For instance, suppose for some $A \in \Sigma, \nu(A) = \rho(A) = \infty$. However, for $a = 1$ and $b = -1$, $\nu - \rho$ is not defined at A as the difference.

(ii) If $a\nu + b\rho$ is defined, the statement follows from Problems 7.22 and 7.24. □

7.27 Let (Ω, Σ) be a measurable space and ξ and η be two complex measures on (Ω, Σ). Show that $\xi + \eta$ is a complex measure on (Ω, Σ) and prove that for any $f \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}) \cap L^1(\Omega, \Sigma; \eta, \overline{\mathbb{R}})$,

$$\int f d(\xi + \eta) = \int f d\xi + \int f d\eta.$$

7.30 Let (Ω, Σ, ν) be a signed measure space and $f \in L^1(\Omega, \Sigma; \nu, \mathbb{C})$. Prove that $f = 0$ ν-a.e. if and only if $\int\limits_A f d\nu = 0$ for all $A \in \Sigma$.

Solution. Let $f = u + iv$. Then, by Definition 7.9 (iv) (in our special case of signed measure ν), $\int f d\nu = \int u d\nu + i \int v d\nu$. This complex integral equals zero if and only if $\int u d\nu = \int v d\nu = 0$. Applying the statement of Problem 7.21 to each integrand we arrive at $u = v = \mathbf{0}$ ν-a.e. The converse of this statement is trivial. □

7.31 Let (Ω, Σ, ξ) be a complex measure space and $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$. Prove that $f = 0$ ξ-a.e. if and only if $\int\limits_A f d\xi = 0$ for all $A \in \Sigma$.

Solution. From Definition 7.9 (iv) we have $\int f d\xi = \int f d\nu_1 + i \int f d\nu_2$. This complex integral equals zero if and only if $\int f d\nu_1 = \int f d\nu_2 = 0$. Applying the statement of Problem 7.21 to f we have $f = \mathbf{0}$ ν_1-a.e. and ν_2-a.e. The converse of this statement is trivial. □

7.33 Show that $(L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}}), +, \cdot, \preceq)$ is a vector lattice over $\mathbb{R}$.

Solution. (i) $L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$ is a vector space. This is due to the fact that for each $f \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$, $\int f d\nu$ is finite and it is a linear functional.

(ii) If $f, g \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$, then so are $f + g$ and $|f - g|$. The latter is due to Problem 7.32. Then, because by Problem 7.17 (i), Chapter 1,

$$f \vee g = \tfrac{1}{2}(f + g + |f - g|)$$

and

$$f \wedge g = \tfrac{1}{2}(f + g - |f - g|),$$

we have $f \vee g, f \wedge g \in L^1(\Omega, \Sigma; \nu, \overline{\mathbb{R}})$. □

7.35 Let $\xi \in \mathcal{C}(\Omega, \Sigma)$ (be a complex measure). Show that a function $f \in L^1(\Omega, \Sigma; \xi, \mathbb{C})$ if and only if $|f| \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}_+)$.

Solution. By Definition 7.9 (iv),

$$f = u + iv \in L^1(\Omega, \Sigma; \xi, \mathbb{C})$$

if and only if $u, v \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}})$. On the other hand, by Problem 7.34, $u, v \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}})$ if and only if $|u|, |v| \in L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}_+)$.

By Definition 7.9(iii), as well as by Problem 7.35,

$$\begin{aligned} &L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}_+) \\ &= L^1(\Omega, \Sigma; \nu_1, \overline{\mathbb{R}}_+) \cap L^1(\Omega, \Sigma; \nu_2, \overline{\mathbb{R}}_+). \end{aligned}$$

Furthermore, by Problem 7.15, $L^1(\Omega, \Sigma, \nu; \overline{\mathbb{R}}) = L^1(\Omega, \Sigma, |\nu|; \overline{\mathbb{R}})$ implying that

$$\begin{aligned} &L^1(\Omega, \Sigma; \xi, \overline{\mathbb{R}}_+) \\ &= L^1(\Omega, \Sigma; |\nu_1|, \overline{\mathbb{R}}_+) \cap L^1(\Omega, \Sigma; |\nu_2|, \overline{\mathbb{R}}_+). \end{aligned}$$

On the other hand, $L^1(\Omega, \Sigma; |\nu_1|, \overline{\mathbb{R}}_+)$ and $L^1(\Omega, \Sigma; |\nu_2|, \overline{\mathbb{R}}_+)$ are semilinear spaces over the semifield $\mathbb{R}_+$, which yields that with

$$|u|, |v| \in L^1(\Omega, \Sigma; |\nu_1|, \overline{\mathbb{R}}_+) \cap L^1(\Omega, \Sigma; |\nu_2|, \overline{\mathbb{R}}_+)$$

we also have

$$|u| + |v| \in L^1(\Omega, \Sigma; |\nu_1|, \overline{\mathbb{R}}_+) \cap L^1(\Omega, \Sigma; |\nu_2|, \overline{\mathbb{R}}_+).$$

Finally, the statement is due to the following inequality:

$$\tfrac{1}{\sqrt{2}}(|u| + |v|) \le |f| \le |u| + |v|. \qquad \square$$

8.1 Prove Proposition 8.2. *Let ν be a signed measure and μ - a positive measure on a measurable space (Ω, Σ). Then the following statements hold.*

(i) $\quad \nu \ll \mu \;\Rightarrow\; \nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow |\nu| \ll \mu$.

(ii) $\nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow \nu \ll \mu$ and $|\nu| \ll \mu$.

(iii) $|\nu| \ll \mu \Rightarrow \nu^+ \ll \mu$ and $\nu^- \ll \mu \Rightarrow \nu \ll \mu$.

Solution. (i) Suppose $\nu \ll \mu$. Let $(P, N) \in [H^+, H^-]_\nu$. By Definition 6.15,

$$\nu^+(E) = \nu(P \cap E) \text{ and } \nu^-(E) = -\nu(P \cap N),\ E \in \Sigma.$$

Suppose E is a μ-null set. Then, by monotonicity of positive measures, $\mu(P \cap E) = \mu(N \cap E) = 0$. Thus, due to absolute continuity of ν,

$$0 = \nu(P \cap E) = \nu^+(E) = \nu^-(E) = |\nu|(E) = 0.$$

Because $|\nu(E)| \leq |\nu(E)|$ we also have $\nu(E) = 0$.

(ii) Suppose $\nu^+ \ll \mu$ and $\nu^- \ll \mu$ and $\mu(E) = 0$. Then, immediately, $|\nu|(E) = 0$ and due to $|\nu(E)| \leq |\nu(E)|$ we also have $\nu(E) = 0$.

(iii) Suppose $|\nu| \ll \mu$ and $\mu(E) = 0$. Then,

$$0 = |\nu|(E) = \nu^+(E) + \nu^-(E)$$

yields

$$\nu^+(E) = \nu^-(E) = 0,$$

also implying that $\nu(E) = 0$. □

8.3 Let (Ω, Σ, ξ) be a complex measure space. Show that $\xi \ll |\xi|$.

Solution. Recall that $|\xi(A)| \leq |\xi|(A)$, for each $A \in \Sigma$, and let $\xi = \nu_1 + i\nu_2$. Suppose $|\xi|(A) = 0$. Then, $|\xi(A)| = 0$ and thus we have

$$|\xi(A)|^2 = (\nu_1(A))^2 + (\nu_2(A))^2 = 0$$

implying that

$$\nu_1(A) = \nu_2(A) = 0 \Rightarrow \nu(A) = 0. \qquad \square$$

8.4 Let $\xi \in \mathcal{C}(\Omega, \Sigma)$. Show that the Radon - Nikodym derivative $\frac{d\xi}{d|\xi|}$ of ξ with respect to $|\xi|$ exists and that the modulus of any Radon - Nikodym density from $\frac{d\xi}{d|\xi|}$ equals $\mathbf{1}$ $|\xi|$-a.e., where $\mathbf{1}$ is the function defined as $\mathbf{1}(\omega) = 1, \omega \in \Omega$.

Solution. Because by Problem 8.3, $\xi \ll |\xi|$, the Radon - Nikodym derivative $\frac{d\xi}{d|\xi|}$ of ξ with respect to $|\xi|$ exists, by the Radon - Nikodym theorem (Theorem 8.9) and it is a subset of $L^1(\Omega, \Sigma; \xi, \mathbb{C})$. This is because $|\xi|$ is finite as the total variation of complex measure ξ. Now, by Problem 7.17,

$$|\xi| = \int \left|\frac{d\xi}{d|\xi|}\right| d|\xi|.$$

Therefore, $\left|\frac{d\xi}{d|\xi|}\right|$ is the Radon - Nikodym derivative of $|\xi|$ with respect to $|\xi|$. Furthermore, because

$$|\xi| = \int \mathbf{1} d|\xi|,$$

it follows that $\mathbf{1} \in \left|\frac{d\xi}{d|\xi|}\right|$. Consequently, all other Radon - Nikodym densities must be equal to $\mathbf{1}$ $|\xi|$-a.e. $\square$

8.6 Let $\xi \in \mathcal{C}(\Omega, \Sigma)$. Show that the Radon - Nikodym derivative $\frac{d\xi}{d|\xi|}$ exists and that for each $g \in L^1(\Omega, \Sigma; \xi, \mathbb{C})$,

$$\int g d\xi = \int g \frac{d\xi}{d|\xi|}\, d|\xi|.$$

Solution. Because by Problem 8.3, $\xi \ll |\xi|$, by Theorem 8.11, given any $g \in L^1(\Omega, \Sigma; \xi, \mathbb{C})$ for any $f \in \frac{d\xi}{d|\xi|}$,

$$\int g d\xi = \int g f\, d|\xi|. \qquad \square$$

8.7 Let μ and ρ be two σ-finite signed measures on a measurable space (Ω, Σ) such that $\mu \approx \rho$ (meaning $\mu \ll \rho$ and $\rho \ll \mu$). Show that

$$\frac{d\rho}{d\mu} = \left(\frac{d\mu}{d\rho}\right)^{-1} \mu\text{- (or } \rho\text{-) a.e.}$$

Solution. We apply the chain rule (Theorem 8.11) replacing ξ in $\xi \ll \mu \ll \rho$ with ρ. Then, we have

$$\frac{d\rho}{d\rho} = \frac{d\rho}{d\mu}\,\frac{d\mu}{d\rho}.$$

On the other hand, $\rho = \int \frac{d\rho}{d\rho} d\rho$ implies that $\mathbf{1} \in \frac{d\rho}{d\rho}$ and therefore,

$$\mathbf{1} \in \frac{d\rho}{d\mu}\,\frac{d\mu}{d\rho}.$$

In other words, for all Radon - Nikodym densities from $\frac{d\rho}{d\mu}$ and $\frac{d\mu}{d\rho}$, there are some whose product equals $\mathbf{1}$. Consequently, the rest of the Radon - Nikodym densities forming the above two classes are such that their respective products equal $\mathbf{1}$ ρ- and therefore μ-a.e. □

8.9 Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space under the following assumptions. Let $\Omega = [0, 1)$, $\Sigma = \mathcal{B} \cap [0, 1)$ and $\mathbb{P} = \mathrm{Res}_{\Sigma}\lambda$ (where λ denotes the Borel - Lebesgue measure). Let $\mathbf{H} = \{H_1, \ldots, H_n\}$ be a finite partition of Ω such that

$$H_k = \left[\tfrac{k-1}{n}, \tfrac{k}{n}\right), k = 1, \ldots, n,$$

and let Σ_0 be the smallest σ-algebra generated by $\mathbf{H}$. Let $X(\omega) = \omega$, for all $\omega \in \Omega$. Calculate a version X_0 of the conditional expectation $\mathbb{E}[X|\Sigma_0]$ and give all other versions of $\mathbb{E}[X|\Sigma_0]$.

Solution. Obviously, we deal here with a special case of Example 8.15, whose results we will utilize throughout this solution. First,

$$\mathbb{P}(H_k) = \tfrac{1}{n}$$

and

$$\int_{H_k} X d\mathbb{P} = \int_{H_k} \omega\, \lambda(d\omega) = \tfrac{2k-1}{2n^2}.$$

Thus, from (8.15c),

$$\mathbb{E}[X|H] = \tfrac{1}{\mathbb{P}(H)} \int\limits_H X d\mathbb{P}$$

and we have

$$\mathbb{E}[X|H_k] = \tfrac{2k-1}{2n^2}\, n = \tfrac{2k-1}{2n}.$$

From (8.15e),

$$X_0 = \textstyle\sum_{n=1}^{\infty} \mathbb{E}[X|H_n] \mathbf{1}_{H_n}.$$

and we have

$$X_0 = \textstyle\sum_{k=1}^{n} \frac{2k-1}{2n}\, \mathbf{1}_{H_k},$$

as a version of the conditional expectation $\mathbb{E}[X|\Sigma_0]$, where

$$\Sigma_0 \ = \sigma(\mathrm{H}) = \{\Omega,\, \emptyset, A : A \text{ is any union of } H_k\text{'s}\}.$$

The same arguments as in Example 8.15(i), lead us to the conclusion that X_0 above is the only version of $\mathbb{E}[X|\Sigma_0]$. □

8.11 Explain exactly why $X \neq \mathbb{E}[X|\Sigma_0]$ if X is not Σ_0-measurable.

Solution. If X is not Σ_0-measurable, then even though the equation

$$\int\limits_A X d\mathbb{P}_0 = \int\limits_A X_0 d\mathbb{P}_0 \quad \text{for all } A \in \Sigma_0$$

holds, the requirement of Corollary 1.21 (that X be Σ_0-measurable) is not met. □

8.13 Prove Proposition 8.17. *Let $X, Y \in \mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, let $\Sigma_0 \subseteq \Sigma$ be a sub-σ-algebra, and $\mathbb{P}_0 = \mathrm{Res}_{\Sigma_0}\mathbb{P}$. Then, for any two real numbers a and b,*

$$\mathbb{E}[aX + bY|\Sigma_0] = a\mathbb{E}[X|\Sigma_0] + b\mathbb{E}[Y|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

Solution. Let $(aX + bY)_0$ be a Radon - Nikodym density of the integral measure $\int (aX + bY)d\mathbb{P}_0$ w.r.t. $\mathbb{P}_0$ and let X_0 and Y_0 be Radon - Nikodym densities of $\int X d\mathbb{P}_0$ and $\int Y d\mathbb{P}_0$, respectively, w.r.t. $\mathbb{P}_0$. Then, we have

$$\int (aX + bY)d\mathbb{P}_0 = \int (aX + bY)_0 d\mathbb{P}_0 \tag{1}$$

as integral measures on Σ_0. Also

$$\int X d\mathbb{P}_0 = \int X_0 d\mathbb{P}_0 \text{ and } \int Y d\mathbb{P}_0 = \int Y_0 d\mathbb{P}_0. \tag{2}$$

By linearity of the integral, the latter two equations (2) yield

$$\int (aX + bY)d\mathbb{P}_0 = \int (aX_0 + bY_0)d\mathbb{P}_0. \tag{3}$$

Comparing (1) and (3) we have

$$\int (aX + bY)_0 d\mathbb{P}_0 = \int (aX_0 + bY_0)d\mathbb{P}_0 \tag{4}$$

good on Σ_0 as measures. As integrals, because both integrands are Σ_0-measurable, (4) implies that $(aX + bY)_0 = aX_0 + bY_0$ $\mathbb{P}_0$-a.s. which exactly is the statement if we turn to the corresponding notation. □

8.15 Prove Proposition 8.19. *Let* $X_1, X_2, \ldots$ *be a monotone nondecreasing sequence of nonnegative r.v.'s from* $\mathbb{L}(\Omega, \Sigma, \mathbb{P}; \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+))$, *let* $\Sigma_0 \subseteq \Sigma$ *be a sub-*σ*-algebra, and* $\mathbb{P}_0 = \text{Res}_{\Sigma_0}\mathbb{P}$. *Then*

$$\lim_{n\to\infty}\mathbb{E}[X_n|\Sigma_0] = \mathbb{E}[\lim_{n\to\infty} X_n|\Sigma_0] \ \mathbb{P}_0\text{-a.s.}$$

Solution. By Radon - Nikodym, for the given sequence $\{X_n\}\uparrow$ there is a sequence $\{X_n^0\}$ of Radon - Nikodym densities, which is $\mathbb{P}_0$-a.s. pairwise monotone, according to Proposition 8.17. There exist point-

wise limits $X = \lim_{n\to\infty} X_n$ and $\widetilde{X}^0 = \lim_{n\to\infty} X_n^0$. On the other hand, applying the monotone convergence theorem (MCT) we have

$$\lim_{n\to\infty} \int X_n d\mathbb{P} = \lim_{n\to\infty} \int X_n d\mathbb{P}_0 = \int X d\mathbb{P}_0 = \nu_0.$$

Thus, there is a Radon - Nikodym density, say X_0, of the integral measure ν_0 relative to $\mathbb{P}_0$ such that

$$\int X d\mathbb{P}_0 = \int X_0 d\mathbb{P}_0 \text{ on } \Sigma_0. \tag{1}$$

We will show that $X_0 = \widetilde{X}^0 \pmod{\mathbb{P}_0}$.

Because X_n^0 is a Radon - Nikodym density of the integral measure $\int X_n d\mathbb{P}_0$ relative to $\mathbb{P}_0$, we have

$$\int X_n d\mathbb{P}_0 = \int X_n^0 d\mathbb{P}_0 \text{ on } \Sigma_0$$

and hence

$$\lim_{n\to\infty} \int_A X_n d\mathbb{P}_0 = \lim_{n\to\infty} \int_A X_n^0 d\mathbb{P}_0, \ \forall A \in \Sigma_0.$$

Then, by the monotone convergence theorem,

$$\int_A X d\mathbb{P}_0 = \int_A \widetilde{X}^0 d\mathbb{P}_0, \ \forall A \in \Sigma_0. \tag{2}$$

Comparing (1) and (2), we have

$$\int_A X_0 d\mathbb{P}_0 = \int_A \widetilde{X}^0 d\mathbb{P}_0, \ \forall A \in \Sigma_0,$$

which yields by Corollary 1.20 that

$$X_0 = \widetilde{X}^0 \pmod{\mathbb{P}_0}.$$

In a nutshell, we have

$$\{X_n^0\} = \{\mathbb{E}[X_n|\Sigma_0]\}$$

and

$$\lim_{n\to\infty}\mathbb{E}[X_n|\Sigma_0] = \widetilde{X}^0.$$

On the other hand, we had

$$\mathbb{E}[\lim_{n\to\infty}X_n|\Sigma_0] = \mathbb{E}[X|\Sigma_0] = X_0$$

and proved that $X_0 = \widetilde{X}^0 \pmod{\mathbb{P}_0}$. □

9.2 Prove that

$$\xi_{a_1}(A_1 \times A_2) = (A_1 \times A_2)_{a_1} = \begin{cases} A_2, & a_1 \in A_1 \\ \emptyset, & a_1 \notin A_1 \end{cases}.$$

Solution. By the definition,

$$\xi_{a_1}(A) = \{\omega_2 \in \Omega_2 : (a_1, y) \in A\}.$$

Thus,

$$\begin{aligned}\xi_{a_1}(A_1 \times A_2) &= \{\omega_2 \in \Omega_2 : a_1 \in A_1, \omega_2 \in A_2\} \\ &= \begin{cases} A_2, & a_1 \in A_1 \\ \emptyset, & a_2 \notin A_1 \end{cases}.\end{aligned}$$ □

9.4 Show the validity of assertion $a)$ in Lemma 9.9. *Let* ξ_{a_i} *be* a_i*-sections for some* $a_i \in \Omega_i$, $i = 1, 2$. *Denote* $\Sigma' = \{A \in \Sigma_\otimes : \xi_{a_1}(A) = A_{a_1} \in \Sigma_2\}$. *Then* Σ' *is a* σ*-algebra in* $\Omega_1 \otimes \Omega_2$.

Solution. (i) $\Omega = \Omega_1 \times \Omega_2 \in \Sigma'$ because, by Problem 9.2, $\xi_{a_1}(\Omega) = \Omega_2$.

(ii) Let $A \in \Sigma'$. Then, $\xi_{a_1}(A) \in \Sigma_2$ and $[\xi_{a_1}(A)]^c \in \Sigma_2$. By Problem 9.3, $[\xi_{a_1}(A)]^c = \xi_{a_1}(A^c)$. Hence, $\xi_{a_1}(A^c) \in \Sigma_2$ and consequently, $A^c \in \Sigma'$. In particular, it follows that $\emptyset \in \Sigma'$.

(iii) We show that Σ' is closed w.r.t. countable unions. Let $\{A_k\} \subseteq \Sigma'$. Then, by Problem 9.3,

$$\xi_{a_1}\left(\bigcup_{k=1}^{\infty} A_k\right) = \bigcup_{k=1}^{\infty} \xi_{a_1}(A_k) \in \Sigma_2. \qquad \square$$

9.6 Show that the Borel - Lebesgue measure λ^n on $\mathcal{B}^n$ and the product measure $\overset{n}{\underset{i=1}{\otimes}} \lambda_i$ on $\mathcal{B}_{\otimes}$ must be identical.

Solution. In Example 9.7 (i) we showed that $\mathcal{B}^n = \mathcal{B}_{\otimes}$. Furthermore, λ^n and $\overset{n}{\underset{i=1}{\otimes}} \lambda_i$ coincide on the semi-ring $\mathcal{S}^n$. Recall that λ^n emerged from elementary content λ^{0n} having the property that

$$\lambda^{0n}((a^n, b^n]) = \prod_{i=1}^{n} (b_i - a_i) = \prod_{i=1}^{n} \lambda_0((a_i, b_i]).$$

The same result for $\overset{n}{\underset{i=1}{\otimes}} \lambda_i$ is due to Theorem 9.11. Now, obviously the premeasure $\widetilde{\lambda}^n$ on the algebra $\mathcal{A}(\mathcal{S}^n)$ and the restriction of $\overset{n}{\underset{i=1}{\otimes}} \lambda_i$ on the same algebra $\mathcal{A}(\mathcal{S}^n)$ must coincide. By Theorem 9.12 (or Corollary 2.14, Chapter 5), the product measure $\overset{n}{\underset{i=1}{\otimes}} \lambda_i$ on $\mathcal{B}^n$ is unique and thus must be equal to λ^n. $\square$

9.8 Let $(\Omega_i, \Sigma_i), i = 1, \ldots, n+k,$ be measurable spaces. Prove that

$$\overset{n+k}{\underset{i=1}{\otimes}} \Sigma_i = \left(\overset{n}{\underset{i=1}{\otimes}} \Sigma_i\right) \otimes \left(\overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i\right).$$

Solution. Obviously,

$$\left(\overset{n}{\underset{i=1}{\otimes}} \Sigma_i\right) \otimes \left(\overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i\right)$$

is generated by the measurable rectangles $A^{(n)} \times B^{(k)}$, where

$$A^{(n)} \in \overset{n}{\underset{i=1}{\otimes}} \Sigma_i \text{ and } B^{(k)} \in \overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i.$$

Thus, the generator $\overset{n+k}{\underset{i=1}{\times}} \Sigma_i$ of $\overset{n+k}{\underset{i=1}{\otimes}} \Sigma_i$ is a subset of the generator

$$\left(\bigotimes_{i=1}^{n} \Sigma_i\right) \times \bigotimes_{i=n+1}^{n+k} \Sigma_i \text{ of } \left(\bigotimes_{i=n+1}^{n+k} \Sigma_i\right) \otimes \Sigma_n.$$

Consequently,

$$\bigotimes_{i=1}^{n+k} \Sigma_i \subseteq \left(\bigotimes_{i=1}^{n} \Sigma_i\right) \otimes \left(\bigotimes_{i=n+1}^{n+k} \Sigma_i\right). \tag{9.8}$$

Let

$$B \in \bigotimes_{i=n+1}^{n+k} \Sigma_i$$

be a fixed measurable set and let

$$\Sigma' = \left\{ A \in \bigotimes_{i=1}^{n} \Sigma_i : A \times B \in \bigotimes_{i=1}^{n+k} \Sigma_i \right\}.$$

Then, Σ' is a σ-algebra that contains $\bigtimes_{i=1}^{n+k} \Sigma_i$. Indeed,

(i) $\bigtimes_{i=1}^{n} \Sigma_i \in \Sigma'$, because a quasi-rectangle

$$A_1 \times \cdots \times A_n \times B \in \bigotimes_{i=1}^{n+k} \Sigma_i.$$

In particular,

$$\prod_{i=1}^{n} \Omega_i \in \Sigma'.$$

(ii) If $A \in \Sigma'$, then $A \in \bigotimes_{i=1}^{n} \Sigma_i$ and thus $A^c \in \bigotimes_{i=1}^{n} \Sigma_i$. Furthermore,

$$\begin{aligned} A^c \times R &= \pi^*_{\{1,\ldots,n\}}(A^c) \cap \pi^*_{\{n+1,\ldots,n+k\}}(B) \\ &= \pi^*_{\{1,\ldots,n\}}\left[\left(\prod_{i=1}^{n} \Omega_i\right) \backslash A\right] \cap \pi^*_{\{n+1,\ldots,n+k\}}(B) \\ &= \left[\pi^*_{\{1,\ldots,n\}}\left(\prod_{i=1}^{n} \Omega_i\right) \backslash \pi^*_{\{1,\ldots,n\}}(A)\right] \cap \pi^*_{\{n+1,\ldots,n+k\}}(B) \end{aligned}$$

$$= (\Omega_1 \times \cdots \times \Omega_n \times B) \backslash (A \times B).$$

The latter is the difference of two sets, with either one belonging to $\overset{n+k}{\underset{i=1}{\otimes}} \Sigma_i$. Thus $A^c \in \Sigma'$.

(iii) Let $S_1, S_2, \ldots \in \Sigma'$. Then, by Problem 5.6$(ii)$, Chapter 1,

$$\left(\bigcup_{i=1}^{\infty} S_i \right) \times B = \bigcup_{i=1}^{\infty} (S_i \times B) \in \overset{n+k}{\underset{i=1}{\otimes}} \Sigma_i,$$

in as much as each $S_i \times B \in \Sigma'$ by the definition of Σ' and the above assumption.

Because $\overset{n}{\underset{i=1}{\times}} \Sigma_i \in \Sigma'$ and $\overset{n}{\underset{i=1}{\times}} \Sigma_i$ is a generator of $\overset{n}{\underset{i=1}{\otimes}} \Sigma_i$, we have the latter be a subset of Σ'. On the other hand, $\Sigma' \subseteq \overset{n}{\underset{i=1}{\otimes}} \Sigma_i$ by the definition of Σ'. Thus, $\Sigma' = \overset{n}{\underset{i=1}{\otimes}} \Sigma_i$.

Now, Σ' was defined w.r.t. a fixed set $B \in \overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i$. Therefore, the same applies to any measurable set B and hence for every $B \in \overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i$ and every $A \in \overset{n}{\underset{i=1}{\otimes}} \Sigma_i$,

$$A \times B \in \overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i.$$

Consequently,

$$\left(\overset{n}{\underset{i=1}{\otimes}} \Sigma_i \right) \times \left(\overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i \right) \subseteq \overset{n+k}{\underset{i=1}{\otimes}} \Sigma_i. \tag{9.8a}$$

Because

$$\left(\overset{n}{\underset{i=1}{\otimes}} \Sigma_i \right) \times \left(\overset{n+k}{\underset{i=n+1}{\otimes}} \Sigma_i \right)$$

is a generator of

$$\left(\bigotimes_{i=1}^{n} \Sigma_i\right) \otimes \left(\bigotimes_{i=n+1}^{n+k} \Sigma_i\right)$$

it follows from (9.8a) that

$$\left(\bigotimes_{i=1}^{n} \Sigma_i\right) \otimes \left(\bigotimes_{i=n+1}^{n+k} \Sigma_i\right) \subseteq \bigotimes_{i=n+1}^{n+k} \Sigma_i. \tag{9.8b}$$

Now, (9.8) and (9.8b) yield that

$$\left(\bigotimes_{i=1}^{n} \Sigma_i\right) \otimes \left(\bigotimes_{i=n+1}^{n+k} \Sigma_i\right) = \bigotimes_{i=n+1}^{n+k} \Sigma_i. \qquad \square$$

9.10 Let $(\Omega_i, \Sigma_i, \mu_i),\ i = 1, 2,$ be measure spaces with σ-finite measures and let $A \in \Sigma_1 \otimes \Sigma_2$. Show that the following statements are equivalent

(i) $\mu_1 \otimes \mu_2(A) = 0,$

(ii) $\mu_2(A_{a_1}) = 0$ μ_1-a.e. on $\Omega_1,$

(iii) $\mu_1(A_{a_2}) = 0$ μ_2-a.e. on $\Omega_2.$

Solution. $(i) \Rightarrow (ii)$. Let $f = \mathbf{1}_A$ and suppose $\mu_1 \otimes \mu_2(A) = 0$. Then $f \in \mathcal{C}_+^{-1}(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$ and by Tonelli's Theorem 9.17,

$$0 = \mu_1 \otimes \mu_2(A) = \int f d\mu_1 \otimes \mu_2$$

$$= \int (\int f_{a_1}(\omega_2) d\mu_2(\omega_2)) d\mu_1(a_1)$$

(due to Problem 9.5)

$$\int (\int \mathbf{1}_{\xi_{a_1}(A)}(\omega_2) d\mu_2(\omega_2)) d\mu_1(a_1)$$

$$= \int \mu_2(\xi_{a_1}(A)) \mu_1(da_1).$$

Because by Lemma 9.10, the function $a_1 \mapsto \mu_2(\xi_{a_1}(A)) \in \mathcal{C}_+^{-1}$, by Lemma 1.15, $\mu_2(\xi_{a_1}(A)) = 0$ μ_1-*a.e.* on Ω_1.

The converse $(ii) \Rightarrow (i)$ is again due to Lemma 1.15.

The proof of $(i) \Leftrightarrow (iii)$ is identical.

Finally, $(i) \Leftrightarrow (iii) \Leftrightarrow (ii)$ yields $(ii) \Leftrightarrow (iii)$. □

9.12 Show that $f_{a_1}^*(A_3) = (f^*(A_3))_{a_1}$, $A_3 \subseteq \Omega_3$.

Solution. We need to show that for each subset $A_3 \subseteq \Omega_3$,

$$(\xi_{a_1} \circ f)^*(A_3) = \xi_{a_1}(f^*(A_3)). \tag{9.12}$$

First notice that ξ_{a_1} does not act like the usual operator. The reader can easily interchange the order on the left of (9.12) to see that it would not make any sense. However, the notation is still more convenient and intuitive. From the definition

$$\xi_{a_1} \circ f = \{(\omega_2, f(a_1, \omega_2)) : \omega_2 \in \Omega_2\}$$

we have

$$(\xi_{a_1} \circ f)^*(A_3) = \{\omega_2 \in \Omega_2 : f(a_1, \omega_2) \in A_3\}. \tag{9.12a}$$

On the other hand, the right-hand side of (9.12), according to the definition of the section operator applied to set $f^*(A_3)$, is

$$\xi_{a_1}(f^*(A_3)) = \{\omega_2 \in \Omega_2 : (a_1, \omega_2) \in f^*(A_3)\}. \tag{9.12b}$$

Now, because maps preserve the inclusion operation, we have

$$f_*\{(a_1, \omega_2)\} \subseteq f_* \circ f^*(A_3) \subseteq A_3. \tag{9.12c}$$

The second inclusion is due to Problem 2.6, Chapter 1. Inclusion (9.12c) is equivalent to $f(a_1, \omega_2) \in A_3$, which makes (9.12b) identical to (9.12a). □

9.14 Show that $\xi_{a_1}\circ(\alpha f+\beta g)=\alpha\xi_{a_1}\circ f+\beta\xi_{a_1}\circ g$.

Solution. We have

$$\xi_{a_1}\circ(\alpha f+\beta g)=\{(y,\alpha f(a_1,y)+\beta g(a_1,y)):y\in Y\}$$

which obviously is a linear combination of two functions from Y to Z and thus the equation holds. □

9.15 Let $f\in\mathcal{C}_+^{-1}(\Omega_1\times\Omega_2,\Sigma_1\otimes\Sigma_2)$ and let a sequence of simple functions $\{s_n\}\subseteq\Psi_+(\Omega_1\times\Omega_2,\mathcal{C}_+^{-1})$ be such that $f=\sup\{s_n\}$. Show that $f_{a_1}=\sup\{(s_n)_{a_1}\}$ (*Hint*: Apply Theorem 6.4, Chapter 5, and Problem 9.14).

Solution. By Theorem 6.4, Chapter 5, f_{a_1} is literally a function of one variable, which, by Proposition 9.15, is Σ_2-Σ_3-measurable and so are functions $(s_n)_{a_1}$. Because f is a pointwise limit of $\{s_n\}$, so obviously is f_{a_1} a pointwise limit of $\{(s_n)_{a_1}\}$. □

9.17 Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be semi-rings on Ω_1 and Ω_2, respectively. Is $\mathcal{G}_1\times\mathcal{G}_2$ also a semi-ring on $\Omega_1\times\Omega_2$?

Solution. The answer is "yes." (i) Let $A=A_1\times A_2\in\mathcal{G}_1$ and $B=B_1\times B_2\in\mathcal{G}_2$. Then, by Problem 5.10 a), Chapter 1,

$$A\cap B=(A_1\cap B_1)\times(A_2\cap B_2),$$

which shows that $\mathcal{G}_1\times\mathcal{G}_2$ is closed w.r.t. finite intersections.

(ii) Let $A=A_1\times A_2\in\mathcal{G}_1$ and $B=B_1\times B_2\in\mathcal{G}_2$. Then, by equation (9.3a) of Lemma 9.3,

$$A\backslash B=[(A_1\backslash B_1)\times A_2]+[(A_1\cap B_1)\times(A_2\backslash B_2)]$$

(because $\mathcal{G}_1$ and $\mathcal{G}_2$ are semi-rings)

$$=(\textstyle\sum_{i=1}^{m}S_i\times A_2)+\left[(A_1\cap B_1)\times\sum_{i=m+1}^{m+n}S_i\right]$$

(using Problems 5.6(ii-iii), Chapter 1)

$$= \sum_{i=1}^{m}(S_i \times A_2) + \sum_{i=m+1}^{m+n}[(A_1 \cap B_1) \times S_i]$$

$$= \sum_{i=1}^{m+n} G_i,$$

where $G_i \in \mathcal{G}_1 \times \mathcal{G}_2$. □

9.19 Let (Ω, Σ, μ) be a σ-finite measure space and let $f \in \mathcal{C}_+^{-1}(\Omega, \Sigma)$. Prove that

$$\int f d\mu = \int_0^\infty \mu(\{f \geq x\})\lambda(dx). \tag{P9.19}$$

(*Hint*: Use Theorem 9.11.)

Solution. We use the change of variables formula in the following form

$$\int g \circ f d\mu = \int g \, d\mu f^*$$

or more specifically

$$\int_\Omega g \circ f(\omega)\mu(d\omega) = \int_{\mathbb{R}} g(x)\mu f^*(dx).$$

In our case, with $g = I$ (identity function) and $f(x) = 0, x < 0$,

$$\int_\Omega f(\omega)\mu(d\omega) = \int_{x=0}^\infty x\mu f^*(dx)$$

$$= \int_{x=0}^\infty \lambda([0,x])\mu f^*(dx).$$

If we define set $A = \{(x,y) \in \mathbb{R}_+^2 : 0 \leq y \leq x\}$, we can identify its x-section $[0,x] = A_x$. So, we then have

$$(\mu f^* \otimes \lambda)(A) = \int_{x=0}^\infty \lambda(A_x)\mu f^*(dx)$$

using the second iterated integral of formula (9.11)

$$= \int_{y=0}^{\infty} \mu f^*(A_y)\lambda(dy),$$

where $A_y = [y, \infty)$. The latter is legitimate, because λ is also σ-finite.

We remark that if $X = f$ is a random variable on a probability space $(\Omega, \Sigma(\Omega), \mathbb{P})$ such that $X \geq 0$ a.s. and $\mathbb{P} = \mu$, then the above formula (P9.19) reads

$$\mathbb{E}X = \int X d\mathbb{P} = \int_{x=0}^{\infty} \mathbb{P}\{X \geq x\}\lambda(dx)$$

and gives an alternative computational formula for the expected value of X. □

9.20 Under the condition of Problem 9.19, let g: $\mathbb{R}_+ \to \mathbb{R}_+$ be a continuous monotone nondecreasing function such that $g(0) = 0$ and which is continuously differentiable on $(0, \infty)$. Show that

$$\int g(f) d\mu = \int\limits_{(0,\infty)} g'(x)\mu(\{f \geq x\})\lambda(dx).$$

Solution. We use the change of variables formula in the following form

$$\int g \circ f d\mu = \int g d\mu f^*$$

or more specifically

$$\int\limits_{\Omega} g \circ f(\omega)\mu(d\omega) = \int\limits_{\mathbb{R}_+} g(x)\mu f^*(dx).$$

Now, we can have

$$g(x) = \int_{t=0}^{x} g'(t)dt = \nu_g([0,x]) \geq 0$$

(because g' continuous and nonnegative) and so

$$\int_{\Omega} f(\omega)\mu(d\omega) = \int_{x=0}^{\infty} \nu_g([0,x])\mu f^*(dx).$$

If we define set $A = \{(x,y) \in \mathbb{R}_+^2 : 0 \leq y \leq x\}$, we can identify its x-section as $[0,x] = A_x$. Obviously, the measure ν_g is σ-finite. So, we then have

$$(\mu f^* \otimes \nu_g)(A) = \int_{x=0}^{\infty} \nu_g(A_x)\mu f^*(dx)$$

using the second iterated integral of formula (9.11)

$$= \int_{y=0}^{\infty} \mu f^*(A_y) d\nu_g(y)$$

and then Theorem 8.10

$$= \int_{y=0}^{\infty} \mu f^*(A_y) g'(y)\lambda(dy),$$

where $A_y = [y, \infty)$. □

Bibliography

[Alb] Albiac, Fernando and Kalton, Nigel J., *Topics in Banach Space Theory*, Springer, New York, 2006.

[Ali] Aliprantis, C.D. and Burkinshaw, *Principles of Real Analysis*, 3rd edition, Academic Press, San Diego, CA, 1998.

[Alt] Alt, Hans W., *Lineare Funktionalanalysis*, Springer-Verlag, Berlin, 1985.

[Ash] Ash, Robert B. and Doléans-Dade, Catherine A., *Probability and Measure Theory*, 2nd edition, Academic Press, San Diego, CA, 2000.

[Aull] Aull, C.E. and Lowen, R. (ed.), *Handbook of General Topology*, Volume 1, Kluwer Academic, Dodrecht, 1997.

[Bart] Bartle, Robert G., *A Modern Theory of Integration*, AMS, Providence, Rhode Island, 2001.

[Bau1] Bauer, Heinz, *Ma□und Integrationstheorie* , Walter de Gruyter, Berlin, 1990.

[Bau2] Bauer, Heinz, *Wahrscheinlichkeitstheorie*, 4th edition, Walter de Gruyter, Berlin, 1991.

[Beh] Behrends, Erhard, *Ma□und Integrationstheorie* , Springer-Verlag, Berlin, 1987.

[Ben] Benedetto, John, J. and Czaja, Wojciech, *Integration and Modern Analysis*, Birkhäuser, Boston, 2009.

[Bou1] Bourbaki, Nicolas, *General Topology*, Part 1, Addison-Wesley, Reading, MA 1966.

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0, © Springer Science+Business Media, LLC 2013

[Bou2] Bourbaki, Nikolas, *Elements of the History of Mathematics*, Springer-Verlag, Berlin, 1994.

[Car1] Carathéodory, Constantin, Über das lineare Maß von Punktmengen - eine Verallgemeinerung des Längebegriffes, *Nachrichten der K. Gesellschaft der Wissenschaften zu Göttingen,* Math.-Phys. Kl. (1914), 404 - 426.

[Car2] Carathéodory, Constantin, *Vorlesungen über Reellen Funktionen*, 3rd edition, Chelsea, New York, 1968.

[Cohe] Cohen, Paul J., *Set Theory and Continuum Hypothesis*, W. A. Benjamin, New York, 1966.

[Cohn] Cohn, Donald, L., *Measure Theory*, Birkhäuser, Boston, 1980.

[Cop] Copson, E.T., *Metric Spaces*, Cambridge University Press, Cambridge, 1968.

[Die] Dieudonné, Jean, *History of Functional Analysis*, North-Holland, Amsterdam, 1983.

[Doob] Doob, Joseph, L., *Measure Theory*, Springer-Verlag, New York, 1994.

[Dsh1] Dshalalow, Jewgeni H., *Real Analysis, An Introduction to the Theory of Real Functions and Integration*, Chapman and Hall/CRC, Boca Raton, FL, 2001.

[Dsh2] Dshalalow, Jewgeni H., *Advances in Abstract Analysis and Applications*, Springer, New York (in preparation).

[Dud] Dudley, Richard M., *Real Analysis and Probability*, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1989.

[Fish] Fisher, Emanuel, *Intermediate Real Analysis*, Springer-Verlag, New York, 1983.

[Fol] Folland, Gerald B., *Real Analysis*, 2nd edition, J. Wiley, New York, 1999.

[Frän] Fränkel, Adolf Abraham, Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre, *Mathematische Annalen*, 86 (1922), 230 - 237.

[Fre] Fremlin, D.H., *Topological Riesz Spaces and Measure Theory*, Cambridge University Press, London, 1974.

[Gil] Giles, J.R., *Introduction to the Analysis of Metric Spaces*, Cambridge University Press, Cambridge, 1987.

[Gor] Gordon, Russell A., *The Integral of Lebesgue, Denjoy, Perron, and Henstock*, AMS, Providence, Rhode Island, 1994.

[Gril] Grillet, Pierre A., *Abstract Algebra*, 2nd edition, Springer, New York, 2007.

[Haw] Hawkins, Thomas, *Lebesgue's Theory of Integration, Its Origins and Development*, Chelsea, New York, 1979.

[Hen] Henstock, Ralph, *The General Theory of Integration, Clarendon Press*, Oxford, 1991.

[Her] Hernández-Lerma, Onésimo and Lasserre, Jean-B., Fatou's lemma and Lebesgue's convergence theorem for measures, *Journ. Appl. Math. Stoch. Anal*., **13**:2 (2000) 137 - 146.

[Hew] Hewitt, Edwin and Stromberg, Karl, *Real and Abstract Analysis*, Springer-Verlag, New York, 1965.

[Hin] Hinderer, Karl, *Grundbegriffe der Wahrscheinlichkeits-theorie*, Springer-Verlag, Berlin, 1972.

[Hrba] Hrbacek, Karel and Jech, Thomas, *Introduction to Set Theory*, Marcel Dekker, New York, 1984.

[Jame] James, I.M. (Ed.), *History of Topology*, Elsevier, Amsterdam, 1999.

[Jech] Jech, Thomas, *Set Theory*, 3rd edition, Springer-Verlag, Berlin, 2006.

[Kel1] Kelley, John L., *General Topology*, Springer-Verlag, New York, 1955.

[Kel2] Kelley, John L. and Srinivasan, T.P., *Measure and Integral*, Volume 1, Springer-Verlag, New York, 1988.

[Kle] Kleiner, Israel, *A History of Abstract Algebra*, Birkhäuser, Boston, 2007.

[Kön] König, Heinz, *Measure and Integration: An Advanced Course in Basic Procedures and Applications*, Springer-Verlag, Berlin, 2009.

[Kun] Kunen, Kenneth and Vaughan, Jerry E. (ed.), *Handbook of Set-Theoretic Topology*, North-Holland, Amsterdam, 1988.

[Kut] Kuttler, Kenneth, L. Modern Analysis, CRC Press, Boca Raton, Florida, 1998.

[LeeP] Lee, Peng Yee and Výborný, Rudolf, *Integral: An Easy Approach after Kurzweil and Henstock*, Cambridge University Press, Cambridge, 2000.

[LeeT] Lee, Tuo Yeong, *Henstock-Kurzweil Integration on Euclidean Spaces*, World Scientific, Singapore, 2011.

[Mac] MacLane, S. and Birkhoff, G., *Algebra*, Chelsea, New York, 1993.

[McD] McDonald, John N. and Weiss, Neil A., *A Course in Real Analysis*, Academic Press, San Diego, CA, 1999.

[Meg] Megginson, Robert E., *An Introduction to Banach Space Theory*, Springer-Verlag, New York, 1998.

[Mich] Michel, H., *Maß und Integrationstheorie* I, VEB Deutscher Verlag der Wissenschaften, Berlin, 1978.

[Mor] Morgan, John C., *Point Set Theory*, Marcel Dekker, New York, 1990.

[Mun] Munkres, James R. *Topology*, 2nd edition, Prentice Hall, Upper Saddle River, NJ, 2000.

[Piet] Pietsch, Albrecht, *History of Banach Spaces and Linear Operators*, Birkhäuser, Boston, 2007.

[Rad] Radon, Johann, Theorie und Anwendungen der absolut additiven Mengenfunktionen, *Sitzungsberichte der Akademie der Wissenschaften in Wien; Mathematischnaturwissenschaftliche Klasse. Abteilung* IIa., 112 (1913), 1295 - 1438.

[Roy] Royden, Halsey and Fitzpatrick, Patrick, *Real Analysis*, 4th edition, Pearson, 2010.

[Rud] Rudin, Walter, *Real and Complex Analysis*, 3rd edition, McGraw-Hill, New York, 1987.

[Ryn] Rynne, Bryan P. and Yongson, Marin A., *Linear Funtional Analysis*, Springer-Verlag, London, 2008.

[Shi] Shirali, Satish and Vasudeva, Harkrishan, *Metric Spaces*, Springer-Verlag, London, 2006.

[Ste] Stein, Elias, M. and Shakarchi, Rami, *Real Analysis, Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, Princeton, NJ, 2005.

[Swa] Swarzt, Charles, *Introduction to Gauge Integrals*, World Scientific, Singapore, 2001.

[Tiet] Tietze, Heinrich, Beiträge zur allgemeinen Topologie I, *Mathematische Annalen*, **88** (1923), 280-312.

[Tych] Tychonov, Andrey, Ein Fixpunktsatz, *Mathematische Annalen*, **111** (1935), 767 - 776.

[Whe] Wheeden, Richard, L. and Zygmund, Antoni, *Measure and Integral*, Marcel Dekker, New York, 1977.

[Wil] Willard, Stephen. *General Topology*, Addison-Wesley, Reading, MA, 1970 (currently by Dover, New York, 2004).

[Yeh] Yeh, J., *Real Analysis, Theory of Measure and Integration*, 2nd edition, World Scientific, Singapore, 2006.

[Zer] Zermelo, E., Untersuchungen über die Grundlagen der Mengenlehre, *Mathematische Annalen*, **65**:2 (1908), 261-281

[Zorn] Zorn, M., A remark on method in transfinite algebra, *Bull. Amer. Math. Soc*., **41**, (1935), 667 - 670.

List of Symbols

Symbol/Notation	Page
$\mathbb{N}$: the set of all positive integers	6
$\mathbb{N}_0$: the set of all nonnegative integers	6
$\mathbb{Z}$: the set of all integers	6
$\mathbb{Q}$: the set of all rational numbers	6
$\mathbb{Q}^c$: the set of all irrational numbers	6
$\mathbb{R}$: the set of all real numbers	6
$\mathbb{C}$: the set of all complex numbers	6
$\mathbb{R}_+$: the set of all nonnegative real numbers	6
$\mathbb{R}_-$: the set of all negative real numbers	6
Union: $A \cup B = \{x \in \Omega\colon\ x \in A \ \vee\ x \in B\}$	6
Intersection: $A \cap B = \{x \in \Omega\colon\ x \in A \ \wedge\ x \in B\}$	6
Disjoint subsets if $A \cap B = \emptyset$	6
Difference: $A \backslash B = \{x \in \Omega\colon x \in A \ \wedge\ x \notin B\}$	7
Symmetric Difference: $A \triangle B = (A \backslash B) \cup (B \backslash A)$	7
Complement with respect to Ω) : $A^c = A^c_\Omega = \Omega \backslash A$	7
": $=$ " reads "set by definition"	7
$\square$ (end of a proof, remarks, examples)	7
$\liminf_{n \to \infty} A_n$ ($\underline{\lim}\ A_n$)	10
$\liminf_{n \to \infty} A_n$ ($\underline{\lim}\ A_n$)	10

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0, © Springer Science+Business Media, LLC 2013

$[X, Y, f]$ 15

$[\mathcal{P}(X), \mathcal{P}(Y), f_*]$ 16

$[\mathcal{P}(R_f),\ \mathcal{P}(D_f),\ f^*]$ 17

$[\mathcal{P}(\mathcal{P}(X)), \mathcal{P}(\mathcal{P}(Y)), f_{**}]$ 17

$[\mathcal{P}(\mathcal{P}(R_f)), \mathcal{P}(\mathcal{P}(X)), f^{**}]$ 17

$[X, Y, I]$ (identity function) 18

$\mathrm{Res}_A f$ (restriction of f to set A) 18

$\preceq$ (a partial order) 29

$[x]_{\mathrm{E}}$ (equivalence class modulo E) 30

$X|\mathrm{E}$ (quotient set of set X modulo E) 30

π_{E} (the projection of set X on its quotient by E) 30

$[X|\mathrm{E}_f, Y, \boldsymbol{f}]$ (reducer of function f) 33

$\prod_{x\in X} Y_x$ (Cartesian product) 43

$\left[\prod_{x\in X} Y_x\,, Y_\alpha\,, \pi_\alpha\right]$ (αth projection map) 44

$\left[\prod_{x\in X} Y_x\,,\ \prod_{\alpha\in A} Y_\alpha\,,\ \pi_A\right]$ (A-projection map) 50

$\pi_A^*(\mathcal{A})$ (an A-cylinder with base $\mathcal{A}$) 50

$\pi^*_{\{\alpha_n,\ldots,\alpha_n\}}(\{f(\alpha_1), \ldots, f(\alpha_n)\}) = \bigcap_{i=1}^{n} \pi^*_{\alpha_i}(f(\alpha_i))$ 51

$\mathrm{card}(A)$ 57

$\aleph_0$ (aleph naught) 59

$\mathfrak{c}$ (cardinal number of continuum) 59

$2^{\mathfrak{c}}$ (cardinal number of hypercontinuum) 62

$\mathbb{F}$ (field) 74

$(X, \mathbb{R},\ +\ ,\ \cdot\ ,\ \preceq\ )$ (partially ordered vector space) 74

$x^+ := x \vee \theta$ (positive part) 77

$x^- := (-x) \vee \theta$ (negative part) 77

$|x|:=x\vee(-x)$ (modulus) 77
Vector spaces:
- l^p space 77
- $\mathcal{C}_{[a,b]}$ space of all continuous functions on $[a,b]$ 78
- $\mathcal{C}^n_{[a,b]}$ space of all n-times differentiable functions on $[a,b]$ 78
- $\mathcal{C}^\infty$ space of all infinitely differentiable functions 78
- $\mathbb{R}^X$ space of all real-valued functions on a set X 78
- $\mathcal{F}_*(X;\mathbb{R})$ all real-valued bounded functions on a set X 78
- $\mathcal{C}(X;\mathbb{R})$ space of all continuous functions 78
- $\mathcal{C}_*(X;\mathbb{R})$ space of all bounded continuous functions 78
- $\mathcal{C}^n(\mathbb{R};\mathbb{R})$ space of all n-times differentiable functions 78
- $\mathcal{C}^\infty(\mathbb{C};\mathbb{C})$ space of all infinitely differentiable functions over $\mathbb{C}$ 78
- $\mathscr{P}$ space of all polynomials with real coefficients 78
- $\mathcal{Q}$ space of all polynomials with rational coefficients 79
span$A$ (linear span) 85
l_c space 88
d (metric) 94
(X,d) (metric space) 94
$(X|\mathrm{E}_d,\boldsymbol{d})$ (reduced metric space) 95
$\mathbb{R}^{\mathbb{N}}$ (real sequences) 96
$d_e(x,y)$ (Euclidean metric) 98
$B(x,r)$ (open ball) 101
$d_\infty(x,y)$ (supremum metric) 101, 102
$\mathring{A}$ or *Int*(A) (interior of a set) 106
$\overline{A}$ or *Cl*(A) (closure of a set) 106
A' (derived set) 107
∂A (boundary of a set) 107

$C(x, r)$ (closed ball) 108
$[\mathbb{N}, X, f]$ (sequence) 112
$\mathcal{C}((X, d); (\mathbb{R}, \rho))$ (continuous functions) 128
$\|\cdot\|$ (norm) 150
$\|\cdot\|_e$ (Euclidean norm) 151
$\|f\|_\infty$ (supremum norm) 153
$(\mathcal{F}_*(X), \|\cdot\|_\infty)$ (space of all bounded functions) 153
$R^1_{[a,b]}$ (Riemann integrable functions) 155
$\left(R^1_{[a,b]}, \|\cdot\|\right)$ (seminormed vector space) 155
$\left(R^1_{[a,b]}|\mathrm{E}_{\|\cdot\|}, \|\cdot\|\right)$ (restrictor-normed vector space) 155
l^∞-space 158
$\|x\|_*$ (standard norm) 167
$(\mathbb{R}, \tau_e)$ (usual topology) 167
$S(0, 1)$ (unit sphere in $\mathbb{F}^n$) 168
τ (topology) 173
(X, τ) (topological space) 173
$(X, \tau(d))$ (metric topology) 174
τ^c (closed sets) 174
$A^\square$ (interior of set A) 175
$\mathrm{Int}(A)$ (interior of set A) 175
$\overline{A}$ (closure of set A) 176
$Cl_X A$ (closure of set A relative to X) 176
∂A (boundary of set A) 177
$N(A)$-tail of $\{x_n\}$ (a tail of $\{x_n\}$ that belongs to set A) 178
$\tau(\mathcal{S})$ (the weakest topology containing collection $\mathcal{S}$) 184
$\tau(\mathcal{S})$ (topology generated by a subbase $\mathcal{S}$) 191

$\mathcal{B} = \overset{n}{\underset{i=1}{\times}} \tau_i$ (base for product topology) 213

$\tau_P = \tau(\mathcal{B})$ (product topology generated by base $\mathcal{B}$) 213

$\overset{n}{\underset{i=1}{\times}} \mathcal{S}_i$ (subbase for τ_P) 213

$\left(\prod_{x \in X} Y_x, \tau_P \right)$ (Tychonov product topology) 216

(Ω, Σ) (measurable space) 241

$\mathcal{B}(\tau)$ (Borel σ-algebra) 247

$\mathcal{B}(\mathbb{R}^n)$ (Borel σ-algebra in Euclidean space) 251

$\mathcal{C}^{-1}(\Omega, \Sigma; \Omega', \Sigma')$ (set of measurable functions) 255

$[\Sigma, \overline{\mathbb{R}}, \mu]$ (set function) 263

(Ω, Σ, μ) (measure space) 264

ε_a (Dirac measure) 265

λ^0 (Lebesgue elementary content) 265

μ_f^0 (Lebesgue - Stieltjes elementary content) 266

$\beta_{n,p}$ (binomial measure) 274

π_α (Poisson measure with parameter α) 274

$\mathfrak{M}(\Omega, \Sigma)$ (set of all measures on a measurable space (Ω, Σ)) 274

$(\mathfrak{M}(\Omega, \Sigma), \mathbb{R}_+, +, \cdot, \preceq)$ (semivector lattice of measures) 277

μ^* (outer measure) 283

Σ^* (σ-algebra of μ^*-measurable sets) 283

$\mathscr{C}_Q(\mathcal{G})$ (countable covers of set Q by "rectangles" from $\mathcal{G}$) 283

$(\mathcal{G}, \gamma)$ (formatter of the outer measure) 284

$\mathcal{N}_\mu$ ($\subseteq \Sigma$ and $N \in \mathcal{N}_\mu \Leftrightarrow \mu(N) = 0$) 290

μ_0^* (restriction of μ^* to Σ^*) 291

λ_c (Lebesgue content on $\mathcal{R}(\mathbb{R}^n)$) 313

λ^* (Lebesgue outer measure) 316

$\lambda_0^* = \mathrm{Res}_{\mathcal{L}^*} \lambda^*$ (Lebesgue measure) 316

$\mathcal{L}^*$ (Lebesgue σ-algebra) 316
$(\mathbb{R}^n, \mathcal{L}^*, \lambda_0^*)$ (Lebesgue measure space) 316
$\lambda = \mathrm{Res}_{\Sigma(\mathcal{S})}\lambda^* = \mathrm{Res}_{\mathcal{B}(\mathbb{R}^n)}\lambda^*$ (Borel - Lebesgue measure) 316
$\mathcal{B} = \Sigma(\tau)$ (Borel σ-algebra) 317
$(\Omega, \mathcal{B}, \mu)$ (Borel measure space) 317
$\mathcal{B}_\mu^*$ (Lebesgue - Stieltjes σ-algebra) 317
$\mu_0^* = \mathrm{Res}_{\mathcal{B}_\mu^*}\mu^*$ (Lebesgue - Stieltjes measure) 317
μ_f^0 (Lebesgue - Stieltjes elementary content) 317
μ_f (Lebesgue - Stieltjes content, premeasure) 317-318
μ_f^* (Lebesgue - Stieltjes measure) 319
μ_f (Borel - Lebesgue - Stieltjes measure) 319
$\mathfrak{B}(\mathbb{R}, \mathcal{B})$ (Borel - Lebesgue - Stieltjes measures) 320
$\mathfrak{D}_e$ (extended distribution functions (d.f.'s)) 321
$\mathfrak{D}_e|\mathrm{E}: = \{\{f; \mu\} : \mu \in \mathfrak{B}\}$ (quotient set of extended d.f.'s) 321
$\mathfrak{B}_*$ (finite Borel - Lebesgue - Stieltjes measures) 321
$\mathfrak{D}$ (set of all d.f.'s) 321
μf^* (image measure) 336
$\mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{R}, \mathcal{B}(\mathbb{R}))$ (real-valued measurable functions) 343
$\mathcal{C}^{-1}(\Omega, \Sigma; \mathbb{C}, \mathcal{B}(\mathbb{C}))$ (complex-valued measurable functions) 343
$\overline{\mathbb{R}}$ (extended real line) 343
$\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ (extended valued measurable functions) 344
$(\overline{\mathbb{R}}^\Omega, \tau_P)$ (extended topology of pointwise convergence) 347
$\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}})|_\mu$ (quotient set modulo μ) 350
$[\mathbf{0}]_\mu$ (set of all measurable functions equal zero μ-a.e.) 350
$[\mathbf{0}]_\mu^+ = [\mathbf{0}]_\mu \cap \mathcal{C}_+^{-1}$ 350
$\mathcal{C}_+^{-1} = \mathcal{C}^{-1}(\Omega, \Sigma, \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+))$ 350
$\Psi_+(\Omega, \Sigma) = \Psi_+(\Omega, \Sigma; \mathbb{R})$ (nonnegative simple functions) 354

$(\mathcal{X}, \mathbb{F}_+)$ (semi-$\mathcal{X}$-space) 355
$\Psi(\Omega, \Sigma) = \Psi(\Omega, \Sigma; \mathbb{R})$ (simple functions) 355
$(\widetilde{\Psi}_+(\Omega, \Sigma), \tau_P)$ (extended nonnegative functions with limits) 355
$f^+ = \sup\{f, \mathbf{0}\}$ 359
$f^- = -\inf\{f, \mathbf{0})$ 359
$L^1(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$-space 370
$\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})$-space 370
$\mathcal{C}_b^{-1}([a, b], \mathcal{B}([a, b]); \mathbb{R})$ (real-valued bounded functions) 399
$L(f,P)$ (lower Darboux sum) 399
$U(f,P)$ (upper Darboux sum) 399
$\mathsf{S}([a, b], \mathcal{B}([a, b]), \mathbb{R})$ (space of all step functions on $[a, b]$) 400
$(\mathrm{R}) \int\limits_{[a,b]} f(x)dx$ (Riemann integral) 401
$(\mathrm{L}) \int\limits_{[a,b]} f(x)dx = \int\limits_{[a,b]} f(x)\lambda(dx)$ (Lebesgue integral) 401
$(L^1([a, b], \mathcal{B}([a, b]), \lambda), \| \cdot \|_{L^1})$ 413
$\nu \ll \mu$ (ν absolutely continuous with respect to μ) 424
$\mathfrak{M}(\Omega, \Sigma)$ is the set of all measures on (Ω, Σ) 427
$\mathfrak{M}_\mu^{\ll} = \{\nu \in \mathfrak{M} : \nu \ll \mu\}$ (measures continuous relative to μ) 427
$\int_\mu$ (mapping defined on $\mathcal{C}^{-1}(\Omega, \Sigma; \overline{\mathbb{R}}_+, \mathcal{B}(\overline{\mathbb{R}}_+))$ 427
$[\mathcal{C}_+^{-1}(\Omega, \Sigma), \mathfrak{M}_\mu^{\ll}, \int_\mu]$ 428
$[\mathcal{C}_+^{-1}(\Omega, \Sigma)|_\mu = \mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}}_+)|_\mu, \mathfrak{M}_\mu^{\ll}, \boldsymbol{\int_\mu}]$ 428
$\frac{d}{d\mu} = \boldsymbol{\int_\mu}$ (reducer of $\int_\mu$) 429
$\frac{d\nu}{d\mu}$ (Radon - Nikodym derivative of ν relative to μ) 429
$f(\alpha, \sigma^2; x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ - \frac{(x-\alpha)^2}{2\sigma^2} \right\}$ (normal pdf) 430
$\mu \perp \rho$ (orthogonal measures) 430
$\mathfrak{P}$, $\mathfrak{N}$ (families of positive and negative sets) 436

(P, N) (Hahn decomposition or Hahn pair) 440
$[H^+, H^-]_\nu$ (Hahn class modulo ν) 440
$\nu^+(\,\cdot\,),\ \nu^-(\,\cdot\,), |\nu| = \nu^+ + \nu^-$ (Jordan variations) 441
$\mathfrak{S}(\Omega, \Sigma)$ (set of all signed measures on (Ω, Σ)) 449
$\mathfrak{S}_*(\Omega, \Sigma)$ (set of all finite signed measures) 449
$(\mathfrak{S}_*(\Omega, \Sigma),\ \preceq\ )$ (vector lattice of finite signed measures) 449
$\|\nu\| = |\nu|(\Omega)$ (norm of signed measure ν) 451
$(\mathfrak{S}_*(\Omega, \Sigma), \|\cdot\|)$ (normed space of finite signed measures) 451
$((\mathfrak{S}_*(\Omega, \Sigma), \mathbb{R},\ +\ ,\ \cdot\ ,\ \preceq\ , \|\cdot\|))$ (Banach lattice) 453
$\mathcal{C}(\Omega, \Sigma)$ (complex measures over the field $\mathbb{C}$) 462
$\|\xi\|$ (norm of a complex measure $\mathcal{C}(\Omega, \Sigma)$) 462
$\mathfrak{S}_\mu^{\ll} = \{\nu \in \mathfrak{S}(\Omega, \Sigma) : \nu \ll \mu\}$ (signed continuous measures) 472
$[\mathbb{L}(\Omega, \Sigma, \mu; \overline{\mathbb{R}})|_\mu, \mathfrak{S}_\mu^{\ll}, \int_{\boldsymbol{\mu}}]$ 472
$\frac{d\nu}{d\mu}, \frac{d\nu^+}{d\mu}, \frac{d\nu^-}{d\mu}$ (derivatives of ν and its Jordan variations) 481
$\mathcal{C}_\mu^{\ll} = \{\xi \in \mathcal{C}(\Omega, \Sigma) : \xi \ll \mu\}$ (complex continuous measures) 483
$\left[L^1(\Omega, \Sigma, \mu; \mathbb{C})|_\mu\,,\ \mathcal{C}_\mu^{\ll}, \int_{\boldsymbol{\mu}}\right]$ 484
$\overset{n}{\underset{i=1}{\times}} \Sigma_i$ (measurable rectangles) 499
$\Sigma_\otimes = \overset{n}{\underset{i=1}{\otimes}} \Sigma_i = \Sigma\left(\overset{n}{\underset{i=1}{\times}} \Sigma_i\right)$ (product σ-algebra) 499
B_{a_i} (a_i-section of set B) 504
$[\Sigma_\otimes, \mathcal{P}(\Omega_2), \sigma_{a_i}]$ (a_i-section of a set) 504
$\overset{n}{\underset{i\,=\,1}{\otimes}} \mu_i$ (product measure) 509
$\varphi_{a_i} \circ f$ (a_i-section of a function f) 513
$\left(\prod_{i=1}^n \Omega_i\,,\ \overset{n}{\underset{i\,=\,1}{\otimes}} \Sigma_i,\ \overset{n}{\underset{i\,=\,1}{\otimes}} \mu_i\right)$ (product measure space) 517
$C_n(x_0\,, r)$ (closed ball in $\mathbb{R}^n$) 517

$\left[\Omega, \prod_{i=1}^{n} \Omega_i, \bigotimes_{i=1}^{n} X_i\right]$ (product map of random variables) 525

$\mathbb{P}_{\otimes X_i} = \mathbb{P}_{\bigotimes_{i=1}^{n} X_i} = \mathbb{P}\left(\bigotimes_{i=1}^{n} X_i\right)^*$ (joint probability distribution) 527

$\left(\prod_{i=1}^{n} \Omega_i, \bigotimes_{i=1}^{n} \Sigma_i, \bigotimes_{i=1}^{n} \mathbb{P}_{X_i}\right)$ (probability product space) 527

Index of Names and Terms

A

αth projection map, 44
αth coordinate, 44
αth factor space, 44
A-projection map, 50
A-cylinder with base $\mathcal{A}$, 50
a.e. property, 349
Abelian algebraic operation, 67
Abelian group, 67
Abelian monoid, 67
Abelian ring, 68
Abelian semigroup, 67
absolutely continuous measure, 424
absolutely continuous signed or complex measure, 471
abstract algebra, 67
accumulation point, 106, 177
additive algebraic operation, 67
additive inverse, 68, 74
additive set function, 263
Al-Khwarismi, Muhammad, 66
$\aleph_0$ (aleph naught), 59
Alexander, James, 172
Alexander subbase theorem, 230
Alexandrov, Pavel, 172
algebra, 66, 240
algebra of functions, 127
algebra over a field, 73
algebraic number, 63
algebraic operation, 67
algebraic structures, 66
associative algebraic operation, 67
associative laws of sets, 7
at most countable set, 59
atomic (discrete) measure, 267
atomic probability measure, 267
axiom of choice, 36
axiom of extent, 7

B

Baire's category theorem, 136
Banach, Stefan, 93, 149, 172
Banach space, 151
base for a topology, 188
base for a topology, criterion 188, 190, 194
base neighborhood, 175
base (simple) rectangle in product topology, 213
base (simple) rectangle, 190
base sets, 188
Beppo Levi's corollary, 383
Bernoulli, Johann, 14
Bernoulli measure, 274

J.H. Dshalalow, *Foundations of Abstract Analysis*,
DOI 10.1007/978-1-4614-5962-0, © Springer Science+Business Media, LLC 2013

bijective (onto and one-to-one) map (function), 17
bilinear transformation, 70
binary relation, 14
binary relation on a set, 29
- antisymmetric, 29
- equivalence, 29
- linear total order, 29
- partial order, 29
- reflexive, 29
- symmetric, 29
- transitive, 29
binomial measure, 274
binomial random variable, 339
Bolzano, Bernard, 3
Bolzano - Weierstrass compactness in metric spaces, 139
Bolzano - Weierstrass compactness in topological spaces, 229-230
Borel, Émile, 141, 262
Borel - measurable bounded functions, 399
Borel measurable function, 257
Borel measure, 317
Borel measure space, 317
Borel outer measure, 317
Borel - Lebesgue measure, 316
Borel - Lebesgue measure of a closed ball in $\mathbb{R}^n$, 517
Borel - Lebesgue - Stieltjes measure, 317
Borel σ-algebra generated by a cocountable topology, 253
Borel σ-algebra, 247
boundary of a set, 107, 177
boundary point, 107
boundary point of a set, 177
bounded function, 134
bounded sequence, 112
bounded set, 34
bounded (d-bounded) set, 122
Bourbaki, Nicolas, 171
box rectangle, 215
box topology, 215
branch of a function, 16

C

$\mathcal{C}^{-1}$-class of functions, 336
canonic chain of partitions, 400
canonic expansion (representation), 355
canonic representation (expansion), 355
Cantor, Georg, 3, 36, 58, 62, 172
Cantor ternary set, 327-328
Cantor's theorem, 61
Cantor's theorem on intersection of closed sets, 133
Carathéodory, Constantin, 240, 285
Carathéodory's extension, uniqueness of, 297, 298
Carathéodory's extension of a formatter, 284, 293

Carathéodory's extension procedure, 283
Carathéodory's extension theorem, 285
cardinal number, 58
cardinality, 58
carrier, 5
Cartan, Henri, 172
Cartesian (direct) product, 14
Cartesian product of an indexed family of sets, 43
Cartesian product of a sequence, 42
Cauchy, Augustin-Louis, 399, 498
Cauchy continuity of a function, 125
Cauchy integrable function, 402
Cauchy sequence, 112
Cauchy sum, 402
Čech, Eduard, 216
chain, 34
chain rule, 487
change of variables for a bijective transformation, 418
change of variables formula, 416
change of variables formula with respect to the affine transformation, 419
characterization of Jordan variations, 444
Chevalley, Claude, 172
choice function, 36
class of complex-valued measurable functions, 343
class of real-valued measurable functions, 343
class of extended real-valued measurable functions, 344
closed ball, 108
closed ball in $\mathbb{R}^n$, 517
closed regular Borel set, 330
closed regular Borel measure, 330
closed set, 174
closed set, 106
closure of a set, 106
closure of a set, 176
closure point for a set, 176
closure point, 106
cluster (accumulation) point, 177
coarser topology, 174
cocountable (countable complement) topology, 178
codomain, 15
cofinite (finite complement) topology, 182
Cohen, Paul, 61
commutative algebraic operation, 67
commutative laws of sets, 7
commutative ring, 68
compact metric space, 138
compact set in topological spaces, 228
compact set under a

continuous function (metric space), 143
compact set under a continuous map in a topological space, 228
compact set, 138
compact sets in Hausdorff spaces, 229
compact topological space, 228
compactness criteria (metric space), 139
compactness criterion (metric space), 140, 145
comparable elements, 29
comparison of sets, 58
complement, 7
complement of set B with respect to set A, 6
complete Carathéodory's extension of a formatter, 284, 293
complete measure space, 290
complete measure, 283
complete metric space, 112
completeness criteria, 132
completion of a measure space, 290
completion of a measure, 290
complex measure space, 458
complex measure, 458
complex vector space, 73
composition of binary relations, 18
composition of continuous functions, 204
composition of maps, 18
composition of measurable functions, 257
conditional expectation given a sub-algebra, 489
conditional probability of an event given a sub-algebra, 490
congruence modulo , 3
congruence, 3
conjugate exponents, 97
consistent extension of a map (function), 18
consistent function, 16
content, 264
continuity at a point, 117
continuity criteria in metric spaces 119, 120
continuity from above at the empty set, 264, 435
continuity from above of a premeasure, criterion of, 272
continuity from above on a σ-algebra, 264
continuity from above on a sequence of sets, 264
continuity from below of a premeasure, criterion of, 271
continuity from below on a σ-algebra, 263
continuity from below on a sequence of sets, 263
continuity of a function, criterion of, 203, 204,

205, 206
continuity of projection maps, 206, 215, 224
continuous from above set functions, 435
continuous from below set functions, 435
continuous function on a dense set, 208, 210
continuous function on a set, 117
continuous function, 203
continuous measure, 424
continuum hypothesis, 61
continuum, 59
convergence in mean, 384
convergence to a point, 178
convergent sequence, 112
coordinate of a function, 44
Coulomb, Jean, 172
countable (denumerable) set, 59
countable complement topology, 178
countably compact topological space, 234
counting measure space, 268
counting measure, 268
cover, 138
criterion of independence of random variables, 527
criterion of Lebesgue integrability, 401

D

d-bounded set, 122
d-bounded function, 134
d_∞-bound of a function, 136
Darboux, Jean-Gaston, 399
De Morgan, Augustus, 66
De Morgan's laws, 8
decomposition of a set, 9
Delsarte, Jean, 172
Denjoy, Arnaud, 406
dense set, 114
dense set, 177
density, 422
denumerable (countable) set, 59
Descartes, René , 42
derived set, 107, 177
diagonal, 200
diameter of a set, 122
difference, 7
dimension of vector space, 87
Dieudonné, Jean, 172
Dirac measure (point mass), 265
direct Riemann integral, 408
direct Riemann integrability, 408
Dirichlet, Peter Gustav, 14, 141
Dirichlet function, 121, 364, 403
Dirichlet's criterion, 411
discrete (atomic) measure, 267
discrete convolution, 71
discrete metric, 95
discrete topology, 173
disjoint family of sets, 9
disjoint sets, 6

distance, 94
distance between two sets, 130
distribution function, 265, 320
distributive laws of sets, 8
domain, 15
Dynkin system, 240

E

Ehresmann, Charles, 172
Eilenberg, Samuel, 172
element of a set, 3
elementary content, 264
elementary event, 5
embedded set, 209
embedding, 209
empty set, 4
equipotent sets, 58
equivalence class modulo $\approx$ (E), 30
equivalence kernel of a function, 31
equivalent classes generated by a function, 31
equivalent metrics, 121
equivalent (similar) norms, 164
Euclidean distance, 98
Euclidean metric, 98
Euclidean norm, 151
Euler, Leonard, 14, 499
event, 5
expectation of a function of a random variable, 419
expectation of a random variable, 371, 418
extended distribution function, 265, 320
extended real line, 343
extended real line topology, 173
extended topology of pointwise convergence, 347
extendibility conditions for the formatter, 284
extendibility of a formatter, criterion of, 294
extendible formatter, 284, 293
extension of the formatter to a measurable space, 284
extension of a map (function), 18

F

F_σ set, 184
factor (quotient) set, 30
factor space, 43
Fatou, Pierre, 361
Fatou's lemma for functions, 384
Fatou's lemma for measures and functions, 393
Fatou's lemma for measures and nonnegative functions, 393
Fatou's lemma for measures, 388
field, 72, 240
figure, 318
finer topology, 174
finite Borel - Lebesgue-Stieltjes measures, 321

finite complement topology, 182
finite dimensional vector space, 87
finite set, 59
finite (ν-finite) set, 435
finite set function, 264
finite signed measure, 435
finite subadditivity, 270
finitely additive set function, 263
first axiom of countability, 179
first countable topological space, 179
formatter of the outer measure, 284, 292
Fourier, Joseph, 14
Fränkel, Abraham, 36
Fréchet, Maurice, 93, 149, 171, 426
Fubini, Guido, 361, 498
Fubini's theorem, 515
function, 14
function continuous at a point, 203
function continuous at a point, criterion of, 203
function f_*, 16
function f_{**}, 17
fundamental system of neighborhoods at a point, 175

G

G_δ set, 184
gamma function, 396
gauge function, 406
gauge integral, 406
generator, 247
geometric measure, 448
Grassman, Hermann, 67
greatest lower bound infimum, 34
Gregory, Duncan 66,
group, 67
automorphism, 69
endomorphism, 69
homomorphism, 68
isomorphism, 68

H

Hahn class modulo, 440
Hahn decomposition, 439
Hahn decomposition theorem, 439
half-open interval (rectangle), 242
Hamel basis for a vector space, 85
Hankel, Hermann, 67
Hausdorff, Felix, 93, 171
Hausdorff space, criterion of, 200
Hausdorff (separated, T_2) topological space 199
Heine, Heinrich Eduard, 120, 141
Heine - Borel property, 161
Heine - Borel theorem, 141
Heine's continuity criteria in metric spaces, 120
Helly, Eduard, 149
Henstock, Ralph, 406

Henstock - Kurzweil integral, 406
hereditary property in topological spaces, 226
Hernández - Lerma, Onésimo, 388
Hilbert, David, 62
Hobson, Ernest, 498
homeomorphic topological spaces, 209
homeomorphism, 209
homeomorphisms and Borel σ-algebras, relationship between, 259
homothetic function, 336
homothetic metric, 149
Hopf, Heinz, 172
Hölder's inequality, 98
hypercontinuum, 62

I

ideal, 278
idempotence, 8
of complement, 8
of intersection, 8
of union, 8
identity function, 18
identity element, 68
image measure, 336
image of a point, 15
image of a set under a function, 16
improper Riemann integral, 406
inconsistent function, 16
indefinite integral, 422
independence criterion of numeric random variables, 352
independence criteria of random variables, 341
independent Dynkin systems of events, 280
independent families of events, 280
independent random variables, 339
independent σ-algebras of events, 280
indicator function, 18
indiscrete topology, 173
infimum, 34
infinite dimensional vector space, 87
infinite set, 59
infinite signed measure, 435
injective map (function), 17
inner regular Borel set, 330
inner regular measure, 330
integrability of a complex-valued function, 462
integral as a linear functional, 371
integral of a complex-valued function with respect to a complex measure, 464
integral of a complex-valued functions with respect to a positive measure, 462
integral of a nonnegative simple function, 362
integral of an extended nonnegative function, 365

integral of an extended real-valued function with respect to a complex measure, 463
integral of an extended real-valued function with respect to a signed measure, 463
integral of an extended real-valued function, 369
integral with respect to Dirac point mass, 366
integration by parts formula for Lebesgue - Stieltjes integrals, 524, 531
interior of a set, 105, 175
interior point, 105, 175
intersection-stable system, 241
intersection, 6
intersection-stable family of sets, 172
into map (function), 17
inverse, 68
inverse image of function f_*, 17
inverse of function , 17
inverse of function f_*, 17
inverse of function f_{**}, 17
invertible map (function), 17
isolated closure point, 106
isometric metric spaces, 338
isometry, 338
isometry between two metric spaces, 124
isomorphic vector spaces, 87
isomorphism, 87

J

join operation, 35
joint probability distribution of random variables, 527
Jordan, Camille, 261
Jordan decomposition of a complex measure, 458
signed measure, 442
Jordan variations, 441

K

kernel, 69
Klein, Felix, 67
Kolmogorov, Andrey, 262, 426
Kronecker. Leopold, 62
Kuratowski, Kazimierz, 172
Kurzweil, Jaroslav, 406

L

l^1 space, 71, 96
L^1 as a vector space, 371
L^1 as a vector lattice, 378
l^∞ space, 158
l^p-norm, 151
l-$\mathcal{X}$-Space, 347
L^1-seminorm, 413
largest element, 34
Lasserre, Jean, 388
lattice, 35
law of double expectation, 490
laws of algebra of sets, 7
least upper bound
supremum, 34
Lebesgue, Henri, 141, 262, 361, 498, 499

Lebesgue content, 313
Lebesgue elementary content, 265
Lebesgue integral, 370
Lebesgue measure, 316
Lebesgue measure space, 316
Lebesgue outer measure, 316
Lebesgue σ-algebra, 316
Lebesgue space, 370
Lebesgue's dominated convergence theorem
for functions, 385, 388
for measures, 389
for measures and functions, 394
for complex-valued functions, 466
Lebesgue's theorem for Riemann integrability, 403
Lebesgue - Stieltjes content, 318
Lebesgue - Stieltjes elementary content, 266, 317
Lebesgue - Stieltjes integral, 370
Lebesgue - Stieltjes measure, 317
Lebesgue - Stieltjes σ-algebra, 317
left distributive law, 68
Leibnitz, Gottfried von, 14
Levi, Beppo, 382, 498
Lie, Marius, 67
limit inferior of a sequence of sets, 10
limit of a sequence, 10
limit point of a sequence, 112, 178
limit superior of a sequence of sets, 10
Lindelöf, Ernst Leonard, 139
Lindelöf compact set in topological spaces, 228
Lindelöf set, 139
Lindelöf space, 139
Lindelöf theorem, 139
Lindelöf topological space, 228
linear (vector) space over a field, 73
linear combination, 85
linear functional, 87
linear hull (span), 85
linear independence, 85
linear manifold, 73
linear map, 87
linear operator, 87
linear span (hull), 85
linearity of the conditional expectation, 493
linearly dependent finite set, 85
linearly disjoint manifolds, 90
Liouville, Joseph, 67
Lipschitz condition, 125
Lipschitz constant, 125
Listing, Johann, 172
lower Baire functions, 402
lower bound, 34
lower Darboux integral, 400

lower Darboux sum, 399
lower limit topology, 195

M

Mandelbrojt, Szolem, 172
map, 15
mapping, 15
marginal sequences, 197
maximal element, 34
maximal linearly
independent set, 89
measurability of a complex-
valued function, 462
measurability of an extended
real-valued function,
criteria of, 345
measurability of a function,
criterion of, 256
measurable
function, 255
rectangle, 499
sets, 241
space, 241
subset, 5
measure, 264
measure generated by an
integral, 422
measure of a hyperplane,
331
measure space, 264
meet operation, 35
mesh of a partition, 400
metric, 94
metric space, 94
metric topology, 173, 174
metrizable topological
space, 174
metrization, 94
minimal element, 34
Minkowski, Herman, 160
Minkowski spaces, 160
Minkowski's inequality, 98
modulus, 77
modulus of congruence, 31
moment generating function,
367
moment generating function
of a random variable, 523
monoid, 67
monotone convergence
theorem
for a sequence of
nonnegative of
functions
on a signed measure
space, 467
for functions, 382
for measures, 392
for the conditional
expectation, 493
generalized, 383
monotone nondecreasing
sequence of sets, 9
monotone nonincreasing
sequence of sets, 9
monotone system, 241
monotone vanishing
sequence of sets, 9
monotonicity
of measures, 268
of outer measure, 283,
285
of the conditional
expectation, 493

Moore, Eliakim, 172
Moore plane, 227
motion-invariant measure, 338
motion, 338
μ-integrable function, 370
μ-minimal decomposition of a set, 270
μ-negligible (negligible) set, 290
μ^*-measurable set, 283, 285
μ^*-separability, 283, 285
μ_0^*-measure, 283
μ-a.e. property, 349
μ-null (null) set, 290
multivalued function, 15
multiplicative algebraic operation, 67
multiplicative inverse, 68

N

$N(A)$-tail of a sequence, 178
$\mathcal{N}_\mu$-set, 290
$N(x,\varepsilon)$-tail, 112
negative part, 77
negligible (μ-negligible) set, 290
negative (ν-negative) set, 435
negative variation of a signed measure, 441
neighborhood base at a point, 175
neighborhood of a function, 51
neighborhood of a point that meets a set, 176
neighborhood of a point, 175
neighborhood system at a point, 175
Nikodym, Otto, 361, 425
non-Borel set, 333
nonnegative simple functions, 354
norm, 150
normal density, 430
normal random variable, 430
normed vector space NVS, 150
nowhere dense set, 114, 177
ν-maximal set, 446
ν-minimal set, 446
null (μ-null) set, 290
null space, 88
NVS, 151

O

one-to-one map (function), 17
onto map (function), 17
open ball with respect to the Euclidean metric, 102
open ball with respect to the supremum metric, 102, 103
open ball, 101
open cover, 138
open cylinders, 194
open map, 222
open neighborhood of function f, 218
open neighborhood of a point, 175

open parallelepiped (rectangle), 190
open rectangle (parallelepiped), 190
open (d-open) set, 103
open sets, 172
open subcover, 138
origin, 73
orthogonal measures, 430
orthogonal transformation, 338
outer measure, 283, 284
outer measure, uniqueness of, 305, 307
outer regular Borel set, 330
outer regular measure, 330

P

pairwise disjoint sets, 9
parallelepiped, 49
partially ordered vector space, 74
partition, 399
partition of an interval, 9
partition of a set, 9
Peacock, George, 66
Peano, Giuseppe, 261
Perron, Oscar, 406
pick-a-point process, 7
point mass (Dirac measure), 265
pointwise convergence, criterion of, 220
Poisson measure, 274
Poisson random variable, 339
positive cone, 74
positive measure, 434
positive part, 77
positive variation of a signed measure, 441
positive (ν-positive) set, 435
Possel, René de, 172
power, 70
power set, 9
pre-image, 17
premeasure, 264
principle of mathematical induction, 37
probability density, 430
probability density function, 430, 489
probability distribution, 338
probability distribution function, 489
probability measure, 264
probability product space, 527
probability space, 264
product map of random variables, 525-526
product measure, 510
product measure space, 517
product measure, uniqueness of, 510
product metric, 99
product σ-algebra, 499
product space, 99
product topology (of two spaces), 197
product topology of arbitrarily many factor topologies, 215
product topology of finitely

many factor topologies, 213
projection map, 44
projection of a set on its quotient, 30
proper subset, 4
proper superset, 4
pseudo-metric, 94
pseudo-metric on a ring, 278
pseudo-metric space, 94

Q

quotient (factor) set, 30
quotient class of extended distribution functions, 321
quotient set modulo μ, 350
quotient topology, 205

R

radius of an open ball, 101
Radon, Johann, 262, 361, 425, 499
Radon measure, 330
Radon - Nikodym density, 422, 429
Radon - Nikodym density of a signed measure, 473
Radon - Nikodym derivative, 429
Radon - Nikodym derivative of a signed measure, 481
Radon - Nikodym derivatives of signed measure and its Jordan variations, 481
Radon - Nikodym theorem, 429
Radon - Nikodym theorem for complex measures, 484
for signed measures, 473
random variable, 338
range, 15
real vector space, 73
rectangle, 49
rectangular cylinder with a base, 49
reducer, 33
reducer of a pseudo-metric, 95
refinement, 399
regular measure, 330
relative topology subspace, 174
restriction of a map (function), 18
restriction of outer measure to Σ^*-algebra, 291
Riemann, Bernhard, 172, 399
Riemann integral, 401
Riemann integrable function, 401
Riemann-Stieltjes integral, 425
Riesz criterion of Heine-Borel property, 162
Riesz, Frigyes, 162, 361, 425
Riesz space, 76
right distributive law, 68
ring, 68, 241
ring generated by a semi-

ring, 250
ring with unity, 68
Russell's paradox, 4

S

Saks, Stanislav, 425
sample space, 5
scalar, 72
Schröder-Bernstein theorem, 62
second axiom of countability, 179
second countable topological space, 179
second separation axiom, 199
section of a function, 513
section of a set, 504
semifield, 73
semigroup of functions, 71
semigroup, 67
semilinear functional, 87
semilinear map, 87
seminorm, 151
seminormed vector space (SNVS), 151
semi-ring, 241
semi-ring, property of, 249
semivector lattice, 77
semivector lattice of measures, 277
semivector space, 73
semi-$\mathcal{X}$-space, 355
separability criterion (metric space), 139
separability of a Lindelöf metric space, 146
separable metric space, 139
separable topological space, 178
separated topological space, 199
sequence, 112
sequence with finite support, 88
sequential compactness (metric space), 139
sequential compactness in topological spaces, 230
sequential limit point of a set, 112, 178
set, 3
set function, 263
set of measures as a lattice, 276
set-algebraic expression, 7
set-algebraic transformation, 7
setwise convergence, criterion of, 391
setwise convergence of measures, 390
setwise limit of measures 390
Sierpiński, Waclaw, 172
σ-additive set function, 263
σ-algebra (-field), 240
σ-algebra generated by
a set, 242, 248
collection of sets, 247
family of functions, 257
function, 255
σ-field, 240
σ-finite set function on a

sequence of sets, 264
σ-finite set function on a system of sets, 264
σ-finite signed measure, 435
σ-finiteness of a set, 297
σ-ideal, 278
Σ-norm, 153
σ-subadditivity, 270
σ-subadditivity of outer measure, 283, 285
Σ^*-σ-algebra, 283
signed Borel measure, 436
signed Borel - Lebesgue-Stieltjes measure, 436
signed measure, 434
signed measures as a Banach lattice, 453
signed measures as a lattice, 448
signed measures as a normed vector space, 451
signed measures as a Riesz space, 450
signed measure generated by the integral, 443
similar (equivalent) norms, 164
simple cylinder with a base, 48
simple functions, 355
simple (base) rectangle, 190
simple (base) rectangle in product topology, 213
single-valued function, 14
singleton, 4
singularity of a measures, 430
smallest element, 34
smallest σ-algebra, 242
SNVS, 151
space $\mathcal{C}^{(n)}_{[a,b]}$ of all n times differentiable functions, 69
sphere, 110
standard norm, 167
standard (natural) topology, 173
step function, 400
Stieltjes, Thomas, 425
stronger finer, larger topology, 174
strongly regular measure, 330
subalgebra, 73
subbase, 191, 192
subbase rectangle, 214
subcover, 138
subpartition of a set, 460
subset, 4
subspace spanned by a set, 85
subspace, 73, 174
subspace of a metric space, 95
superset, 4
supremum, 34
supremum metric, 96, 101, 102, 134
supremum norm, 153
surjective map (function), 17
symmetric difference, 7
systems of sets, diagram of, 244

T
τ-closed sets, 174
T_2 space, 199
Tietze, Heinrich, 215
TIH metric, 150
Tonelli, Leonida, 498
Tonelli's theorem, 514
topological space 173
topology, 172
topology generated by a
  metric, 120
  subbase, 191
  set, 173
topology of pointwise
  convergence, 220
total variation of a complex
  measure, 459
total variation of a signed
  measure, 441
totally bounded set, 122
transformation, 15
translation, 337
translation-invariance of
  Borel - Lebesgue mea-
  sure, 339
translation-invariant metric,
  149
translation-invariant Borel
  measure, 323
Trennungsaxiom, 199
triangle inequality, 94
Tukey, John, 37
Tukey's lemma, 37
two-points compactification,
  343
two-sided identity, 67
Tychonov, Andrei, 216
Tychonov's theorem, 232
Tychonov topology, 216

U
Umgebung, 175
unbounded set, 122
uncountable set, 59
uniform continuity criterion
  in compact space, 144
uniform convergence, 153
uniform convergence
  of continuous bounded
  functions, 135
uniform (supremum) metric,
  134
uniform random variable,
  339
uniformly bounded set of
  functions, 136
uniformly continuous
  function, 123
union, 6
unit cylinder with a base, 46
unity function, 18
unity, 68
universe, 5
upper Baire function, 402
upper bound, 34
upper Darboux integral, 400
upper Darboux sum, 399
Urysohn, Pavel, 172
usual topology, 167, 173

V
vaguely hereditary property
  in topological spaces, 226
Valée-Poussin de la, Charles,

498
vector, 73
vector lattice, 76
vector (linear) space over a field, 73
Vitali, Giuseppe, 262

W

weak topology generated by a family of functions, 219
weaker coarser, smaller topology, 174
weakest topology containing a family of sets, 184
weakly hereditary property in topological spaces, 226
weakly regular measure, 330
Weierstrass, Karl, 172
well-ordering principle, 36
well-ordered set, 35
Weil, André, 172
Weyl, Hermann, 172

X

$\mathcal{X}$-space, 347

Z

Zermelo, Ernst, 3, 36
Zermelo's theorem, 36
zero, 68
Zorn, Max, 37
Zorn's lemma, 37